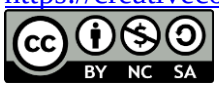
UiT The Arctic University of Norway

Faculty of Humanities, Social Sciences and Education

Searching for a framework to analyze critical thinking and rational argumentation in online educational discussions

Jens Breivik

A dissertation for the degree of philosophiae doctor – May 2020



Year: 2020

Title: Searching for a framework to analyze critical thinking and rational argumentation in online educational discussions

Name: Jens Breivik

Print: UiT The Arctic University of Norway

Preface and acknowledgements

I am completing my thesis for submission after months of the lockdown of the university to stop the spread of the Corona virus. The lockdown of schools has actualized how digitalization is a key issue for education. The option of using digital technology for interaction among learners, in addition to the dissemination of learning material, is more urgent than ever. I am sure that future research on how digital technology impacts education will be quite different than projects initialized before this situation.

Deciding what to believe and to do – on a rational basis – is one conception of what critical thinking and rational argumentation is about. This has become even more important during these last months of confusion and insecurity. Critical thinking is an educational virtue. Still, it is important to discuss what we understand by critical thinking.

Since the department of education is moving into a new building on campus as I am preparing the thesis for submission, I am packing down my office: books, piles of paper and other reminiscences of last years' work life. This reminds me of my initial motivation when I started the Ph.D. program: to articulate a vision of education that emphasizes students' curiosity and drive for learning. Rather than pushing students to have to say or write something to fulfill mandatory requirements, I believed in stimulating students to have something to say because it is important or urgent for them. The distinction between *having to say something* and *having something to say* fascinated me. Further, I believed digital media enables immense access to information and new options for student interaction. How could university education make the most of these opportunities to stimulate students' curiosity and engagement? Later, the project turned into a narrower one, scrutinizing a tiny piece in this larger puzzle.

The anamnesis of the partly forgotten process of writing this Ph.D. thesis is fascinating. Glancing through piles of scribbled notes, printed articles and commented-upon drafts, I am puzzled by all the good pieces of advice I have received and the efforts of colleagues discussing with me, reading drafts, commenting and suggesting. Unfortunately, I have not been able to pursue more than a fraction of the good pieces of advice and ideas. I am grateful for all the inputs and help I have received and the opportunity to spend so much time on this work

Professor Mariann Solberg has been my supervisor through the whole process of developing this Ph.D. thesis. Thank you, Mariann, for your encouragement, patience, joyful discussions and constructive feedback to both developed drafts and newborn thoughts. Having you as an ally has been of great value to me.

Professor Marit Allern was my co-supervisor for the earlier stages of this Ph.D. thesis. Thank you, Marit, for your support, care, and challenges!

My Ph.D. project was funded by UiT The Arctic University of Norway and located at the group for university pedagogics, which later merged into Result: Centre for Teaching, Learning and Technology. A warm thank you to my former colleagues there for a good working environment. Providing pedagogical support and training for the academic staff of the whole university is among the most inspiring teaching experiences I ever have had. Meeting inspired and skillful educators from a broad range of disciplines has been stimulating. One thing I have learned is that faculty development balances on a fine line between, on the one hand, scholarly and research-based groundings and, on the other hand, practical applications.

A great number of other people, friends and colleagues have contributed to my work and motivated me in diverse ways. I am grateful to all of you. One of those I have discussed the most diverse area of topics with is my friend and colleague through several years, Pål Anders Opdal. Thank you for thought-provoking ideas and constructive comments to drafts throughout the process.

Being a part of NATED, the National Graduate School in Educational Research and meeting experienced scholars – Professor Monika Nerland, Professor Peter Maassen and Professor Andreas Lund – and inspired Ph.D. candidates was a great pleasure and was important for finding my way through this process. Warm thanks to all of you! A special thank you to Kristine Ludvigsen for reading and providing important feedback to drafts of articles and the extended abstract.

About a year ago, I started in a new position at the department for education. There, I met welcoming and inspiring colleagues. I am happy to be a part of the collegium! A special thank you to Head of Studies Ellen Karoline Dahl, you have been extraordinary in motivating me as well as arranging a work schedule that has made it possible for me to complete this work. The research group in the philosophy of education and our lively discussions has also been an inspiring environment. Thanks to all of you!

Professor Tove Irene Dahl was appointed as the midway evaluator. Thank you for your inspiring and supportive pieces of advice. I am sure that if I had followed your ideas, this thesis would have been better and maybe completed before now. Last spring, Professor Andreas Lund read the three articles and a draft of the extended abstract thoroughly. Thank you for acknowledging comments, both constructive and critical! Based on your feedback, I have written a much better extended abstract.

To friends and family, I am grateful that you are the people you are. Some of you have been interested in my Ph.D. work, the topics and the progress, and some of you not. I have found both approaches inspiring and relieving.

Anton, Johan and Sigurd – you are the three favorite persons in my life, the ones I most of all enjoy discussing with and hanging around with in the mountains. Thank you for being so patient and generous with your father!

Jens Breivik

Tromsø, May 2020

Summary

Teachers set up online discussion boards to promote student interaction thereby enhancing students' ability to think critically and present rational arguments. Researchers seek to investigate how participation in online discussions affects students learning and how the facilitation of educational discussion influences student participation in such discussions. To do this, researchers have suggested several frameworks and coding schemes to analyze such discussions. In this thesis, I evaluate and discuss two such frameworks and their coding schemes designed to analyze online discussions.

One of my main claims is that the framework that has dominated the research field – the Community of Inquiry framework and its associated Cognitive Presence construct – fails when analyzing and assessing students' critical thinking and rational argumentation in online discussions. The Community of Inquiry framework builds on the idea that progress through different discussion phases towards a solution phase is a good indicator of critical thinking. The proponents of the framework base this on John Dewey's idea that human thoughts are responses to problems or obstacles, and thus thinking is a way to handle – occurring problems. I claim that even though Dewey's naturalistic and pragmatic approach to thinking provides an interesting perspective on thinking, progress through phases is a weak indicator of the quality of thinking. Discussion and collaborative thinking may contain a high level of critical thinking and rational argumentation without reaching a solution. Similarly, discussion and collaborative thinking may reach a solution even if the thinking lacks a critical nerve, and the arguments are weak.

Another acknowledged framework for analyzing critical thinking and rational argumentation in online discussions that I discuss builds on Toulmin's argument model. This model focuses on arguments' microstructure and emphasizes how discussants back their claims and limit subsequent extensions. Applied as a coding scheme on discussion posts, this model uses the occurrence of backed claims and claims that are sufficiently limited as indicators of high quality argumentation. Yet there are problems related to this approach to analyze and score discussions.

In my discussion of these two frameworks to analyze critical thinking and rational argumentation in online educational discussions, I focus on: 1) What philosophical ideas form the basis for the frameworks? How are these ideas operationalized into a coding scheme with observable indicators? 2) Do the discussed frameworks adequately operationalize critical thinking, rational argumentation, and collaborative knowledge building? Focusing on

adequacy, I discuss whether a coding scheme and its observable indicators make it possible to draw valid inferences about a theoretical construct.

Three articles and an extended abstract comprise the thesis. Study 1 is a discussion of the Cognitive Presence construct in the Community of Inquiry framework. In Study 2, I use the categories arguments' microstructure (based on Toulmin's argument model) and argumentations macro-structure to analyze transcripts from online discussions in an introductory philosophy course. The aim is twofold: first, to understand how students construct arguments, and second, to discuss the adequacy of categories of arguments' micro- and macro-structure as a means of analyzing discussions. Study 3 gives an overview of how researchers use frameworks based on Toulmin's argument model to analyze online educational discussions. Further, the study discusses the adequacy of such frameworks. In the extended abstract, I situate the research in a broader context; account for methodological approaches to this kind of theoretical inquiry; and discuss implications, responses to the work, and limitations of the study.

Sammendrag (Summary in Norwegian)

Undervisere tar i bruk diskusjonsforum for å fremme kommunikasjon mellom studenter og dermed styrke studentenes evne til å tenke kritisk og presentere rasjonelle argumenter. Forskere forsøker å finne ut hvordan deltakelse i nettbaserte diskusjoner påvirker studentenes læring og hvordan underviseres tilrettelegging påvirker studentenes deltakelse i nettbaserte undervisningsdiskusjoner. For å gjøre dette har forskere utviklet en rekke rammeverk og kodingsskjema for å analysere slike diskusjoner. Formålet med denne avhandlinga er å evaluere og diskutere to rammeverk som er utviklet for å analysere nettdiskusjoner, og deres kodingsskjema.

En av mine hovedkonklusjoner er at rammeverket som har dominert forskingsfeltet, det såkalte *Community of Inquriy*-rammeverket, og konstruktet *Cognitive Presence* fra dette rammeverket, mislykkes i å analysere og vurdere studentenes kritiske tenkning og rasjonelle argumentasjon i nettdiskusjoner. *Community of Inquiry*-rammeverket bygger på ideen at progresjon gjennom ulike faser av en diskusjon mot en løsning, er en god indikator for kritisk tenkning. Forskerne som etablerte rammeverket baserer dette på Deweys ide om at menneskelig tenkning er responser på problemer eller hindringer, og at tenkning dermed kan forstås som en respons på – og dermed vår måte å løse eller håndtere – problemer vi støter på. Jeg argumenterer for at selv om Deweys naturalistiske og pragmatiske redegjørelse for tenkning er et interessant perspektiv, så er progresjon gjennom ulike faser en dårlig indikator for kvalitet på tenkning. Diskusjoner og tenkning i fellesskap kan holde høyt nivå når det gjelder kritisk tenkning og rasjonell argumentasjon, og likevel ikke føre til en løsning. Tilsvarende kan diskusjoner og tenkning i fellesskap føre til en løsning selv om tenkning er ukritisk og argumentene er svake.

Et annet anerkjent rammeverk for å analysere kritisk tenkning og rasjonell argumentasjon i nettdiskusjoner bygger på Toulmins argumentmodell. Jeg diskuterer dette også dette rammeverket i avhandlinga. Denne modellen belyser argumenters mikrostruktur, med vekt på hvordan diskusjonsdeltakere begrunner og angir begrensninger for sine påstander. Når denne modellen danner grunnlaget for et kodingsskjema, er det forekomst av begrunnede påstander, og forekomst av argumenters begrensning, som fungerer som indikatorer for kvalitet på tenkning og argumentasjon. Det er problemer også med denne måten å analysere og vurdere diskusjoner.

I min diskusjon av disse to rammeverkene som er utviklet for å analysere kritisk tenkning og rasjonell argumentasjon i nettbaserte undervisningsdiskusjoner reiser jeg følende

spørsmål: 1) Hva slags filosofisk ide danner grunnlaget for rammeverkene? Hvordan er disse ideene operasjonalisert i kodingsskjema med observerbare indikatorer? 2) Operasjonaliserer disse rammeverkene kritisk tenkning, rasjonell argumentasjon og kollektiv kunnskapsutvikling på en adekvat måte? Ved å fokusere på om rammeverk og indikatorer er adekvate, diskuterer jeg om disse forskningsverktøyene danner grunnlag for å tekke valide slutninger

Avhandlinga består av tre artikler og kappe. Artikkel 1 er en diskusjon av Cognitive Presence-konstruktet fra Community of Inquiry-rammeverket. I artikkel 2 analyserer jeg utskrifter fra nettdiskusjoner i et innføringskurs i filosofi ved hjelp av kategorier fra argumenters mikrostruktur (basert på Toulmins argumentmodell) og makrostruktur i argumentasjon. Formålet med denne analysen er todelt: For det første, å kartlegge hvordan studenter konstruerte argumenter, og for det andre, å diskutere hvor adekvate kategoriene argumenters mikrostruktur og makrostruktur i argumentasjon er som verktøy for å analysere slike diskusjoner. I artikkel 3 gir jeg en oversikt over hvordan forskere har brukt Toulmins argumentmodell i analyser av nettbaserte undervisningsdiskusjoner. Deretter diskuterer jeg hvor adekvate rammeverk basert på denne modellen er for analyser av nettdiskusjoner. I kappa lokaliserer jeg denne forskninga i en breiere kontekst og gjør rede for metodologiske spørsmål knyttet til denne typen teoretiske undersøkelser. Videre diskuterer jeg implikasjoner, respons på arbeidet og begrensinger i avhandlinga.

Table of Contents

List of enclosed articles

Study 1

Breivik, J. (2016) Critical thinking in online educational discussions measured as progress through inquiry phases: A discussion of the cognitive presence construct in the community of inquiry framework. *International Journal of E-Learning & Distance Education/Revue internationale du e-learning et la formation à distance, 31*(1).

Study 2

Breivik, J. (2020a) Argumentative patterns in students' online discussions in an introductory philosophy course: Micro- and macrostructures of argumentation as analytic tools. *Nordic Journal of Digital Literacy, Vol. 15* (No. 1-2020), 8-23.

Study 3

Breivik, J. (2020b) Toulmin's argument model used to analyze critical thinking in online educational discussions: An overview and critical evaluation. Under review by *Computers & Education*[1].

[1] I submitted a previous version of the manuscript to Computers & Education in October 2019, and I received an invitation to consider major revisions and resubmit. The enclosed version was revised and resubmitted to Computers & Education in May 2020.

List of figures

List of tables

1. Introduction

In this thesis, I discuss how to analyze critical thinking in online educational discussions. I will critically examine two established coding frameworks used in research on online educational discussions. The frameworks I discuss are 1) the Community of Inquiry model (Garrison, Anderson, & Archer, 1999) and the Cognitive Presence construct from this model, and 2) frameworks that build on Toulmin's argument model (see for instance Cho & Jonassen, 2002; Clark & Sampson, 2008; Weinberger & Fischer, 2006). I will present these frameworks below and in the articles.

Critical thinking and rational argumentation are ideals of higher education. Online educational discussions are used to enhance such competencies. An overarching research problem is:

a) How does participation in online educational discussions enhance students' capacity for critical thinking and rational argumentation?

This question forms the background of my work, and I address the topic quite briefly in one of the studies; however, it opens up another research problem:

b) How can students' critical thinking and rational argumentation be analyzed in online educational discussions?

My thesis is centered around this challenge, which I split into two research questions:

1. How is critical thinking analyzed in the Community of Inquiry framework, and how adequate is this operationalization?
2. How do frameworks building on Toulmin's argument model analyze quality of argumentation, and how adequate is this way of operationalizing rational argumentation?

Overall, the aim is to critically evaluate these frameworks and thereby contribute in a discussion about research tools and methods in this specific research field on online, educational discussions.

I will comment upon how I address the research questions in the specific articles later in this introduction, after I have presented the research context.

1.1. Background: Frameworks used to analyze online educational discussions

Online educational discussions (hereafter abbreviated to OED) are a commonly used learning activity[2]. Educators and researchers strive to explore how participation in OEDs influences students learning. Further they strive to understand how different strategies for facilitating OEDs, such as role-play, challenges to solve genuine problems, or requests for structured arguments affect how students contribute. To investigate these topics, researchers need tools to analyze the quality of discussions.

Among the intended outcomes of educational discussions are enhanced ability to critical thinking and rational argumentation. Evaluating tenability of claims, providing rational backing for arguments, and thinking critically make discussions an arena for learning subject matter. Many researchers of OEDs have thus focused on critical thinking and rational argumentation.

Several reviews (De Wever et al., 2006; Martono & Salam, 2017; Wise & Paulus, 2016) have pointed out that the most frequently used method to analyze discussion transcripts from OEDs is content analysis, based on a predefined coding scheme. Nevertheless, content analysis includes a number of approaches and is not a standardized technique (Gerbic & Stacey, 2005; Neuendorf, 2002). Other methods used in OED research are thematic analysis (transcript analysis based on grounded and inductive approaches), social network analysis (how relations between discussants are established and maintained), pre- and post-tests that aim to identify learning acquisition, and surveys or interviews of students' and/or teachers' perceptions of learning in OED.

The research community has suggested a vast number of frameworks and coding schemes for content analysis of transcripts from OED. However, no consensus has been established on their adequacy analyzing online educational discussions. Two review articles (De Wever et al., 2006; Weltzer-Ward, 2011) sum up challenges related to validity and choosing an appropriate framework for research. In a frequently cited review article, De Wever et al. (2006) discussed several coding schemes that researchers have proposed to

[2] In the thesis, I use the term Online Educational Discussion (OED). Other terms used in the research field are Asynchronous Online Discussion (AOD) – see, for instance, De Wever, Schellens, Valcke, and Van Keer (2006) or Schindler and Burkholder Jr (2014); Computer Mediated Communication (CMC) – see Naidu and Järvelä (2006); Computer Conferencing, see Garrison et al. (1999); Asynchronous Learning Environment; and "e-tivities," (Salmon, 2002).

analyze OED transcripts. They pointed out a need to raise "questions about the coherence between theoretical base and the operational translations of the theory in the instruments. Instruments are hardly compared or contrasted with one another. [….T]he validity of the instruments [is] limited (De Wever et al., 2006, p. 6)." Their study did not draw any conclusions about coding schemes, except to point to the general need to compare frameworks and discuss theoretical backgrounds. In a review on research on OED, Weltzer-Ward (2011) identified more than 50 different coding schemes published before April 2010. Still, researchers suggest new coding schemes (see Biasutti (2017). Weltzer-Ward aimed to provide a basis for a more consistent use of coding schemes and facilitate comparison between studies based on different schemes. In her article, she suggested that the research community should attempt to reach an agreement on one or a couple of the most frequently used coding schemes in the field. Weltzer-Ward (2011) did not dig into discussions about the theoretical basis for or validity of the different schemes, except to assume that the research community has validated the most frequently used schemes.

In this thesis, I take up the challenge suggested by De Wever et al. (2006) and actualized by Weltzer-Ward (2011); namely, the need to scrutinize the adequacy and theoretical basis of two commonly used coding schemes. This is in line with Schindler and Burkholder Jr (2014)'s conclusion on their review of research on online discussions as a means to stimulate critical thinking: "Researchers should identify a clear definition of critical thinking and a comprehensive description of how critical thinking should be demonstrated and assessed in AODs" (Schindler & Burkholder Jr, 2014, p. 24).

In chapter 2, I will present a brief overview of research on digital tools in education and how OED may contribute to student learning.

1.2. Frameworks to be discussed

The frameworks I have selected for further discussion are suitable for such discussion for several reasons. Both are among the most frequently used frameworks and coding schemes in the research field. According to Wise and Paulus' (2016) overview of the research field, "argumentation [models] and knowledge development models [are] the most common models used for investigating learning in online discussions" (Wise & Paulus, 2016, p. 274). The two frameworks I have selected represent each of these models. Correspondingly, Weltzer-Ward (2011) found in her review that the Community of Inquiry-model was "widespread accepted" (p. 56) and "dominat[ed] the literature" (p. 67). Additionally,

"variations on Toulmin's (1958) argument framework which provides a theoretical basis for describing argument construction have also been extensively employed" (p. 69).

Both frameworks are inspired by established and acknowledged philosophical theories, yet their approaches to operationalize critical thinking are quite different. This makes them suitable for comparison.

Community of Inquiry and the Cognitive Presence Construct

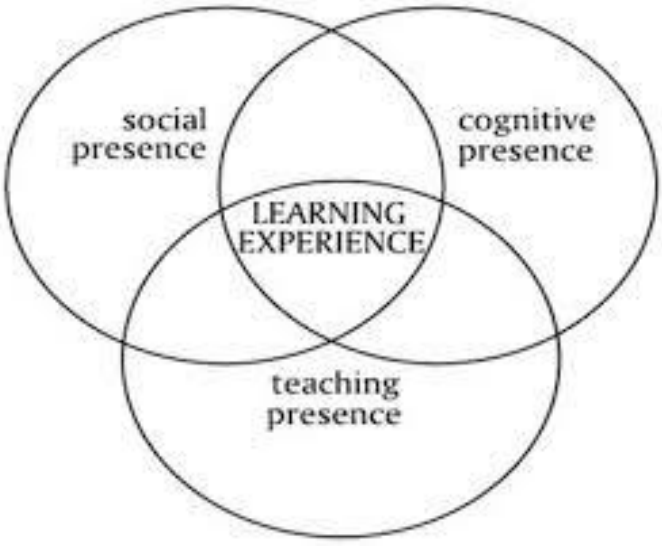
The Community of Inquiry framework was developed at the University of Alberta around the millennium as a theoretical framework to promote and research e-learning understood as a "convergence of technological and pedagogical developments". A key idea is that digital technology provides opportunities to merge *independence* and *interaction* and thus that "students should be actively engaged in the sustainable communities of inquiry" (Garrison, 2011, p. 1). A primary focus is OED as media to foster such educational rationales.

Inspirations for the framework are John Dewey's (Dewey, 1920, 1933/1986, 2007) thoughts about education and knowledge as social enterprises and his idea that inquiry, and thinking in general, is a response to the obstacles of human projects.

The Community of Inquiry model (Figure 1) suggests the following three distinct but

Figure 1

The Community of Inquiry model. Originally published in Garrison et al. (1999).



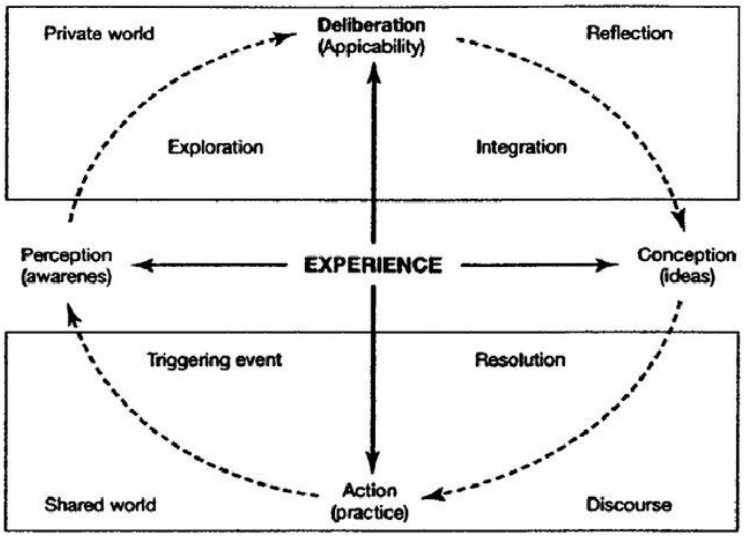
overlapping constructs to assess online educational interaction: social presence, teaching presence, and cognitive presence. Social presence signifies participants' ability to present

themselves as “real people” in a purely textual medium and is characterized by emotional expression, open communication, and group cohesion. Teaching presence describes the design and facilitation of the educational experience. Cognitive presence concerns how learners construct meaning in a critical community of inquiry (Garrison, Anderson, & Archer, 2001).

I will focus on the Cognitive Presence construct, which aims to describe higher-order knowledge acquisition and critical thinking. The Practical Inquiry model (Figure 2) illustrates how Cognitive Presence during inquiry processes combines the dimensions of *action* to *deliberation* and *perception* to *conception*. Further, the model illustrates how practical inquiry includes four phases (*triggering event*, *exploration*, *integration*, and *resolution*) that describe cognitive presence” (Garrison, 2011, p. 46).

Figure 2

The Practical Inquiry model. From Garrison (2011).



A coding scheme was developed by Garrison et al. (1999, 2001) to assess cognitive presence and thereby critical thinking. The central assumption is that progress through the phases of inquiry processes serves as an indicator for the quality of critical thinking.
The four inquiry phases (in ascending order: *triggering event*, *exploration phase*, *integration phase*, and *resolution phase*) form the basis for the coding scheme, which enables interpretation and categorization of OED messages to detect which inquiry phase and level of critical thinking they represent. According to this coding scheme, all phases of an inquiry process are required to achieve high levels of critical thinking. A discussion that reaches a tested solution represents a more advanced level of critical thinking than a discussion that

hovers around identifying or understanding a problem. Crucial is the idea that "deep and meaningful learning does not occur until students move to integration and resolution stages" (Shea et al., 2010, p. 15).

In Study 1, my discussion of the model addresses the idea that reaching the final integration and resolution phases can be viewed as an indicator of higher levels of critical thinking and thereby create deep and meaningful learning.

In addition to the coding scheme for content analysis of discussion transcripts, researchers have developed a survey instrument based on the same constructs and indicators (Arbaugh et al., 2008). The survey instrument, too, uses progress through phases of inquiry processes as indicator of the cognitive presence construct, and thereby critical thinking. The discussion of the adequacy of progress towards the ultimate resolution phases is relevant to the survey instrument as well.

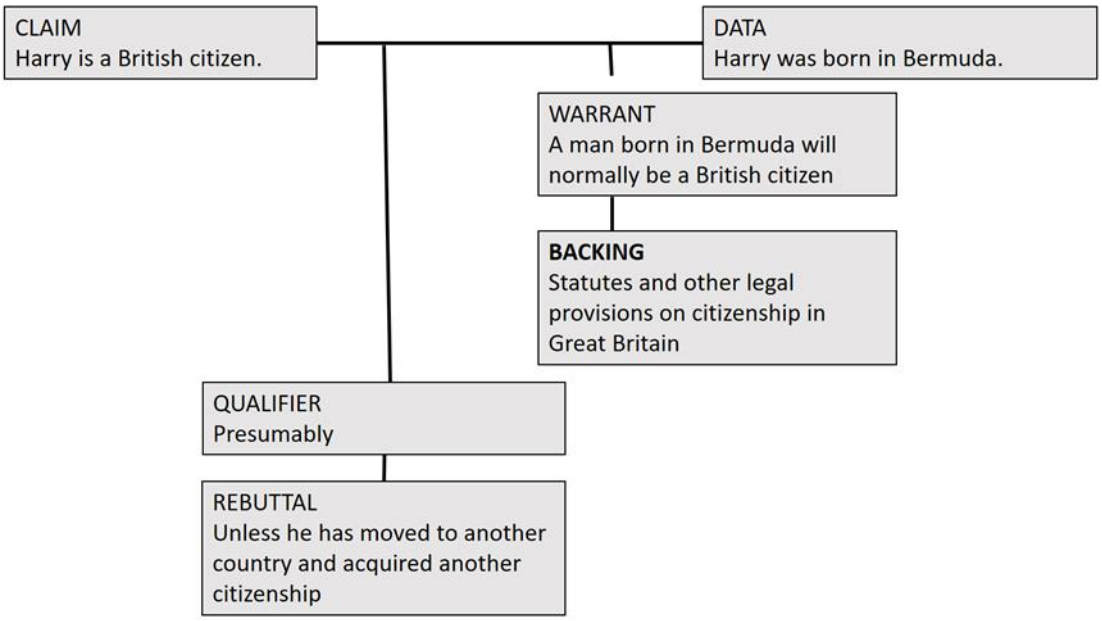
Frameworks building on Toulmin's argument model

Another group of commonly used frameworks for analysis of OEDs is based on Toulmin's argument model (Toulmin, 2003/1958; Toulmin, Rieke & Janik, 1984). Central to these frameworks is the idea that the presence of certain argument components serves as an indicator of argument quality, and this is what I take up for discussion.

According to Toulmin's model, an argument consists of advancing a *claim*, a proposition that the utterer holds true and wants to defend. The claim needs support in order to be justified and accepted as true. *Data* work as evidence for claims. In many cases, we will need something that bridges data to the claim and guarantees that the data really provide evidence for the claim. Toulmin's *backing* is the premise that guarantees the inference from data to evidence. Backing may need some kind of support, which this model names the *warrant*. The data, warrant, and backing may support a claim with varying degrees of certainty. The claim may follow with necessity or a certain degree of probability. The estimate of the probability of the claim, known as the *qualifier*, is part of the line of argument as well. The circumstances under which the claim is not true is the rebuttal. Figure 3 illustrates the model.

Figure 3

Toulmin's argument model.



Several researchers have used the model to analyze argumentation in OED. In Study 2, I demonstrate how a simplified version of this model is commonly used for analysis of OEDs. In Study 3, I present an overview of how the Toulmin model is used in the research field. Further, I sketch Toulmin's background for his work and discuss how adequate the presence of argument components indicates argument quality.

1.3. Relationship between research questions and the three studies

In this thesis, I will critically evaluate two ways of analyzing students' contributions to OEDs. In one of them, progress through certain phases of inquiry is the key indicator of critical thinking. The other uses the presence of certain argument components as the key indicator of argument quality.

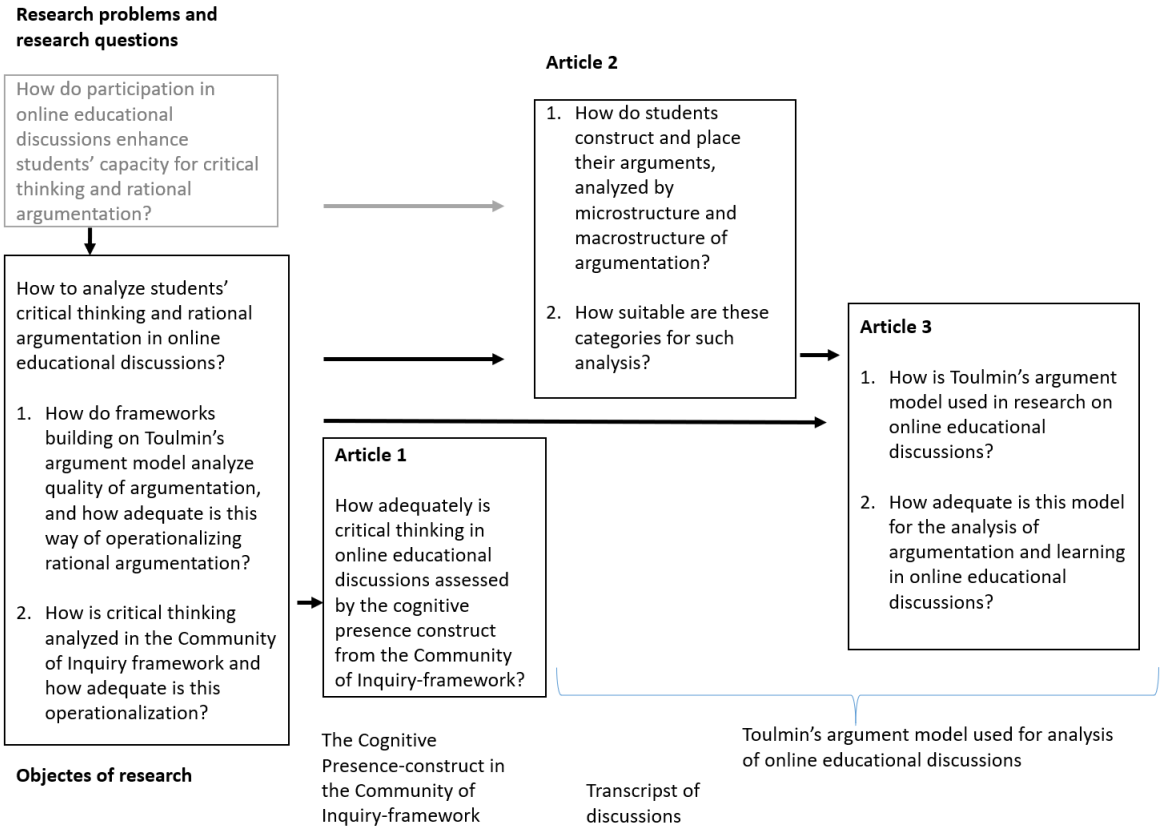
Figure 4 illustrates the relationship between an overarching research problem in this field and my research questions in the articles. The background research problem is marked in grey, as I only to a limited degree pursue this in one of the articles and in the thesis.

In the first study, "Critical thinking in online educational discussions measured as progress through inquiry phases: A critical discussion of the cognitive presence construct in the community of inquiry framework" (Breivik, 2016), I raise the question about this framework's construct validity and what kind of empirical findings the framework has

generated. My search for an adequate framework to analyze transcripts from OEDs led me to scrutinize one of the most acknowledged and frequently used models in the field. My conclusion in the study is that this way of operationalizing critical thinking is dubious.

Figure 4

Research problems and research questions in the thesis.



In the first study, "Critical thinking in online educational discussions measured as progress through inquiry phases: A critical discussion of the cognitive presence construct in the community of inquiry framework" (Breivik, 2016), I raise the question about this framework's construct validity and what kind of empirical findings the framework has generated. My search for an adequate framework to analyze transcripts from OEDs led me to scrutinize one of the most acknowledged and frequently used models in the field. My conclusion in the study is that this way of operationalizing critical thinking is dubious.

In the second study "Argumentative patterns in students' online discussions in an introductory philosophy course - Micro- and macro structures of argumentation as analytic tools" (Breivik, 2020), I use another set of established categories of analysis used for research on online educational discussions, namely argumentation' macro-structure and arguments'

micro-structures. These are based on Toulmin's argument model. By analyzing discussion transcripts, Study 2 shows the potentials and pitfalls of OEDs as a learning activity. Even if reviewers (Weltzer-Ward, 2011; Wise & Paulus, 2016) find frameworks based on Toulmin's model both widely accepted and extensively used, my analysis questions the adequacy of such frameworks.

In Study 3, "Toulmin's argument model used to analyze critical thinking in online educational discussions: An overview and critical evaluation", I present an overview and the constraints of how researchers have used (and continue to use) Toulmin's model as an analytic tool for OED. My aim is to follow up on Study 2 and the second research question.

In chapter 4, I sum up my conclusions from the three studies as well as the answers to the research questions.

1.4. Key concepts in the research questions

Before describing how my research project was developed, I will comment upon some key words and concepts I use in phrasing the research questions. These are highlighted below:

1. How is *critical thinking* analyzed in the Community of Inquiry *framework* and how *adequate* is this operationalization?
2. How do *frameworks* building on Toulmin's argument model *analyze* argumentation quality, and how *adequate* is this *operationalization*?

Critical thinking and rational argumentation are key terms in the research questions and my thesis. These concepts also serve as ideals for higher education. No clear baseline is established on how to understand and operationalize critical thinking, rational argumentation and the relationship between them.[3] Conceptualizing critical thinking and rational argumentation is a theme in my thesis, but its relevance extends beyond the context of OEDs. I have commented upon critical thinking and rational argumentation in a separate chapter in both the extended abstract and the articles. I maintain that evaluating the tenability of claims is a core aspect of both critical thinking and rational argumentation. This is in line with several scholars in the field. I return to this topic in chapter 2.

[3] See, for instance Siegel (1988/2013): "[T]he notion of critical thinking remains obscure and ill-defined; the theoretical conflicts between the various analyses offered to date are significant," or Davies and Barnett (2015): "[A]fter more than four decades of scholarly work, critical thinking remains more elusive than ever." I return to this topic in chapter 2.

The objects for my study are *frameworks and operationalizations*, also designated as coding schemes. What is a framework precisely? In research literature on content analysis of OEDs, the word framework is commonly used to describe several approaches to evaluate discussion quality. Different frameworks may focus on different aspects of quality like critical thinking or educationally valuable talk (Uzuner, 2007). Thus, the term framework designates something more specific than an overarching theory (for instance constructivism, a cognitive or socio-cultural perspective on learning). Researchers need to operationalize frameworks addressing one or another aspect of discussion quality into a coding scheme. This consist of identifying observable indicators that represent the quality under investigation (here: critical thinking and rational argumentation). The term coding and coding scheme is associated with quantitative content analysis, which is the most common variant of analysis employed in the research I comment upon. Despite this, content analysis of discussion transcripts may be performed qualitatively as well. Thus, I have used the terms framework and operationalization to designate how the quality under investigation (here: critical thinking and rational argumentation) are described, and what is held to be observable indicators.

I use the term *analyze* in the research questions to describe what researcher use with frameworks/operationalizations/coding schemes for. Alternative terms might be *measure* or *assess*. The term *measure* is related to quantitative analysis; again, I have used a more general term to signal that the question I discuss is relevant in both quantitative and qualitative analysis. Assessments are commonly used to evaluate student's work and provide feedback. *Define* is another relevant term. To define something implies identifying necessary and/or sufficient properties. Educators may use the frameworks and operationalizations I discuss for assessment and feedback, among other purposes, although the purpose of my focus is the analysis of discussion transcripts for research.

Adequate is another key term in the research questions. What I aim to ask is: do the frameworks/coding schemes analyze or measure what they are intended to? Another term for this might be *construct validity* (based on Messick, 1995; Shadish, Cook, & Campbell, 2002). In a recent guidebook (AERA, NCME, & APA, 2014, p. 11), the concept of validity is reserved for descriptions of the quality of inferences that can be drawn, not the quality of instruments or data.[4] Traditions and scholars employ concepts related to validity differently.

[4] This contrasts how the concept of validity is used in the quote from De Wever et al. (2006, p. 6) in section 1.1: "the validity of the instruments are limited" (p.6). Here, validity is used to describe a quality of an instrument rather than the inferences that use of an instrument may enable.

Rather than digging into such discussions about different conceptions of validity in the introduction, I have used the term *adequate* when formulating the research questions. In Study 1, I based my analysis and discussion on the term *construct validity* and the related terms *construct relevance* and *construct representativeness.* I chose not to use these terms in the following article, since *construct validity* is used and defined in contrary ways. I will return to the concept of validity in the method section. When reading the term adequate, one may ask, "Adequate for what purpose?" In my wording of the research questions, and throughout the thesis, I use the general concept of adequacy to address whether frameworks and coding schemes serve to identify general qualities of critical thinking and rational argumentation.

1.5. Impetus for and development of the thesis

My PhD project was initiated and funded by a program at UiT – The Arctic University of Norway that aimed to increase the use of ICT to provide more flexible education. The project was located in a center for faculty development where Scholarship of Teaching and Learning (SoTL) was a central research approach (Boyer, 1990). My original aim was to research how digitalization in higher education would enhance *Bildung* (Solberg, 2011). I chose OED to study this, since educational discussion boards provide opportunities to utilize digital tools to increase interaction among students and thereby provide an arena to foster generic competencies.

At the time when the project was initiated, the university emphasized the importance of distance education. The Faculty for Humanities, Social Science and Education established introductory courses in several subjects, and discussion boards embedded in the learning management system were supposed to make up for distance and to enhance student interaction in off-campus programs. Experience showed that the activities in and quality of online discussions varied. As the only mode of peer interaction, the discussion boards were a vital part of instructional design, and educators sought knowledge on how to best facilitate such discussions. Some of my colleagues developed a model to describe the quality of students' interaction in OED (Anfinsen & Laugerud, 2007). This model focused on dynamics (level of interaction among discussants); level of academic argumentation (discussants' use of knowledge and subject terms from their studies instead of personal experiences or opinions); and level of reflection (discussants' justification of claims and ability to compare different kinds of justification). I recognized that this topic deserved further development.

My ambition was to conduct research that combined my interest in the digitalization of higher education (I previously occupied a position in an agency digitalizing higher education funded by the Ministry of Education[5]) with my interest for dialog-oriented education. I had also previously worked with Philosophy for Children (P4C) as a teacher and researcher (see for instance Breivik & Løkke, 2007). I planned to put my philosophy education to use and further educate myself as an empirical researcher.

I soon found the concept of *Bildung* problematic. The concept *Bildung* has a prominent history in European and Scandinavian education. A few years before I started my project, a so-called *Bildung* committee released their report on *Bildung* in Norwegian higher education (Dannelsesutvalget, 2009). It appeared overly ambiguous to me. Certain instances of *Bildung* describe education as it seeks to inaugurate individuals into a cultural canon. Others allude to the critique of utilitarianism in education or the development of autonomous judgment; or generic skills. *Bildung* signifies rather different, even contradictory, educational ideals. In my approach, I found the ability to think critically and argue rationally more suitable competencies for research. I left the concept of *Bildung* and focused on critical thinking.

One of the first steps in my project was to search for a suitable theoretical framework for research on online educational discussions focusing on critical thinking. Among the first I arrived at was the Community of Inquiry framework. The phrase "community of inquiry" was derived from John Dewey and James Peirce and is well established in the tradition of Philosophy for Children. Matthew Lipman, a leading proponent for philosophy as a pedagogical approach in school, viewed the communities of inquiry as essential (Lipman, 2003; Lipman, Sharp, & Oscanyan, 1980).

I recognized in the existing literature an extensive number of other approaches to analyzing learning in general and critical thinking in OEDs. It soon became clear to me that frameworks in the research field deserved to be theoretically and methodologically scrutinized. The abundance of frameworks and the lack of scholarly debate and consensus about their adequacy lit a spark in me. Unclear core concepts in empirical research invite conceptual and philosophical approaches.

In their review on organizational studies, Alvesson and Sandberg (Alvesson & Sandberg, 2011; Sandberg & Alvesson, 2011) suggested that in the social sciences,

[5] Former *Norgesuniversitetet – Norway Opening Universities*. This agency has since been merged into the larger directorate *Diku*.

researchers generate research questions in two complementary ways – gap spotting or problematizing. Gap spotting involves reviewing existing fields of research and identifying under-researched knowledge fields yet to be filled. Problematizing involves challenging underlying assumptions of existing research. In my thesis, I apply both approaches. Previous reviews (se for instance De Wever et al., 2006) have identified confusion in the research field and a need to scrutinize the appropriateness and validity of suggested frameworks, which I address.

In the beginning, my plan was to conduct one study discussing theoretical frameworks. However, this intriguing topic cannibalized most of my PhD project. My initial aspiration to conduct empirical research to inform design of online courses led to the broader discussion of theoretical and methodological assumptions. The joy of theoretical discussion isolated from educational practice was not my impetus or motive for the thesis, but rather a method to focus on problems within empirical and practical contexts. By doing this I hope to bridge theoretical and philosophical matters and the practical application of abstract concepts like critical thinking in empirical research. The aim is not theoretical discussions for the sake of theory; rather, my theoretical discussions aim to support the improvement of empirical research. Better research tools may contribute to better research and in turn better educational practices. Although my focus is on empirical research methods, my discussion on the operationalization of critical thinking and rational argumentation may have implications for educational practice. Not all implied readers of this work are developers or users of the research tools I discuss. Yet my discussion of the research tools hopefully contributes to awareness of the complexities of concepts like critical thinking and rational argumentation in educational settings.

Initially, higher education was the context for my study, and I planned to do empirical research on the use of OED in universities. As my focus shifted towards research tools, as much as the empirical phenomenon, the context of higher education context became less significant. The research literature I addresses focuses to a large extent on higher education settings, yet not exclusively (Loncar, Barrett, & Liu, 2014). Some studies researched OED in high school settings. Critical thinking and rational argumentation have served as educational ideals since the rise of academic culture in the antiquity and is actualized in the attention to 21st century skills. This attention towards generic competencies and deep learning have also influenced primary and secondary education. In a Norwegian context the education reform

Fagfornyelsen (The Subject Renewal) a significant example (Kunnskapsdepartementet, 2016; NOU 2015:8, 2015).

1.6. Structure of the thesis

In this introduction, I have presented the overall research problem, my rationale, and the research questions I pursued in the three studies. Further, I have delineated the scope of the thesis and introduced the frameworks under evaluation. This displays how the thesis, including the three studies, are situated in the research field and form an integrated whole.

In chapter 2, I sketch the landscapes wherein the discussions occur, critical thinking and rational argumentation as educational ideals, and the digitalization of education.

In chapter 3, the methods chapter, I present generativity – building on previous research as an overall research quality. I also introduce aspects of validity, and coherence in research design as a strategy to enhance validity. Coherence in research design concerns alignment between research questions, methods, and inferences. Some of the research questions I pose in the thesis and the three studies are conceptual, and I present the methods I used to approach them. Coherence also concerns alignment between philosophy of science-assumptions, research questions, and methods. I present a pragmatist philosophy of science that dismisses the schism between (post-)positivism and constructivism in favor of endorsing diverse methodology. Finally, I discuss ethical considerations concerning the research community, informants, and society in general. Since the quality of research instruments is an overall concern, the methods chapter forms a substantial part of this extended abstract.

In chapter 4, I sum up how I approached the research problems and questions in the three studies. My answers differ from the current consensus in the field. The frameworks I evaluate are considered to be validated and important in the research field. Even though, I conclude that they lack adequacy as general operationalizations of overall qualities of critical thinking and rational argumentation.

In chapter 5, I view the thesis from an alternate perspective. My overall aim was to engage in a scholarly debate: Have I succeeded? Based on responses to Study 1, I comment on how I have succeeded in this. Further, I comment on topics that I have omitted from my discussions, and sketch relevance for wider contexts. My own evaluation of the inferences I draw are apparent throughout chapters 3, 4 and 5, where I discuss the coherence in, and limitations of, my study.

2. Broader landscape

The research motive in my thesis questions how to analyze critical thinking. I discuss this in the context of research on OED. In the introduction, I described how the question of how to analyze critical thinking is notably urgent in research on OED. The topic is actual outside the context of OED as well. Existing literature is rich on questions concerning critical thinking and rational argumentation as educational ideals. In this chapter, I will briefly sketch some themes from this literature as a background, yet the discussions in the three studies in my thesis focus on how these topics are addressed in OED research.

Digitalization of education have been of great interest to educators, researchers, and policy makers over last decades, and this theme pertains to the questions I discuss in my thesis. Even if I address vital questions within research about digitalization of education, the three studies do only to a limited degree dig into potentials of digitalization of education. Rather, digitalization and research on OEDs serve as an arena to discuss questions about how to operationalize critical thinking and rational argumentation. Nevertheless, in Study 2, I also displayed potentials and pitfalls of using OEDs as a learning activity.

2.1. Critical thinking and rational argumentation

In this section, I establish some issues concerning critical thinking and rational argumentation as educational rationales to provide insight on how these capacities may be analyzed in research on OEDs. I hope to demonstrate that the topic of my thesis serves as a lens to ponder over crucial questions in this landscape.

The ability to think critically and argue rationally have been embraced as ideals for higher education since the dawn of academic institutions in antiquity. Interest for such educational objectives in primary and secondary education has increased over the past decades as well through the use of terms like "21st century skills," "generic competencies," and "graduate attributes." Due to the ubiquitous access to information through digital channels, the emphasis in education has moved from the *possession* of knowledge to the *processing* of knowledge.

Despite the wide acceptance of critical thinking and rational argumentation as learning objectives, teachers commonly complain about students' lack of progress in these areas. The research literature echoes this concern. Arum and Roksa (2011) demonstrated in their book *Academically Adrift* that large cohorts of American college students showed weak or no improvement in their skills in critical thinking and complex reasoning during their two first

years of college (up to 45% of the students), and even after the completion of four years of college (36% of the students). Other studies have demonstrated similar reasons for concern regarding how students develop skills in critical thinking (Cahill & Bloch-Schulman, 2012).

One possible reason for the lack of improvement may be found in teachers' instructional techniques (or lack thereof) to facilitate critical thinking. Paul, Elder, and Bartell (1997) found that 89% of the instructors in their study of Californian teacher educators claimed that developing students' critical thinking was an important objective for them. However, only 19% of the instructors were able to demonstrate a clear conception of critical thinking, and as little as 9% actually taught critical thought (see also Inch & Warnick, 2011, p. 9).

The lack of a clear conception of critical thinking is not unique for the teachers in the study by Paul et al. (1997). Rather, it is characteristic of the field. In an introduction to a comprehensive handbook on critical thinking in higher education, Davies and Barnett (2015) claimed that "[a]fter more than four decades of scholarly work, critical thinking remains more elusive than ever" (Davies & Barnett, 2015, p. 3). This echoes Barnett's (1997) previous observation that "[h]igher education […] which prides itself on critical thought has done no adequate thinking about critical thinking." Yet another important scholar in the field, Siegel (Siegel, 1988/2013) argued that "the notion of critical thinking remains obscure and ill-defined; the theoretical conflicts between the various analyses offered to date are significant."

Moore (2011b) demonstrated the confusion about the concept of critical thinking and how to convey that idea to students among a sample of faculty members from humanities. By analyzing their conceptions of critical thinking, he identified seven (partly overlapping and conflicting) conceptions of critical thinking:

- Judgement – ability to evaluate and take a stand towards presented claims
- Fallibilism – a skeptical and provisional view of knowledge
- Originality – seeing new connections, creativity
- Ability to read texts on their own premises
- Rationality – assessing the validity of inferences
- Ethical awareness and an activist stance towards contemporary problems
- Self-reflexivity – ability to question own assumptions

These seven conceptions illustrate how faculty members understood critical thinking quite differently. This complexity may lead to confusion for students and teachers who must analyze students' contributions, for instance in online discussions.

Several possibly divergent reasons to promote critical thinking add to the confusion. Some proponents of critical thinking argue that the ability to think critically, present rational arguments, and evaluate the tenability of arguments are taken to be at the very core of every academic and rational endeavor for knowledge and understanding (see for instance Davies & Barnett, 2015). In line with this, the connection between deep learning (Biggs & Tang, 2007; Marton & Säljö, 1976) and higher order learning (Anderson & Krathwohl, 2001) reinforces the ability to provide and evaluate justifications for claims, not only memorize and reproduce them. Other scholars posit critical thinking as a means of problem solving and innovation vital for innovation and knowledge-driven economies. Poce, Corcione, and Iovine (2012) and the OECD (2010) have developed this line of argumentation. Another approach to the societal need for critical thinking highlights the value of enlightened and reflective citizens as a prerequisite for a functioning democracy. Dewey (2007) and Habermas (1987) are proponents of this line of reasoning. Finally, some scholars, pointing to a more existential motive, see the ability to think critically as essential for individuals to act as independent beings (Bailin & Siegel, 2002; Siegel, 1988/2013). According to this view, ability to provide reasons and rational justifications for one's choices essential are essential for being treated as an autonomous being.

Despite several reasons to promote critical thinking as educational virtues, objections may be raised. Critical thinking may lead to disruption in the form of decreased respect for traditions and authorities. Siegel (1988/2013) labels such objections as the *ideology objection* and the *indoctrination objection*. The ideology objection claims that critical thinking is an ideal that cannot, like any other ideology, be justified alone. The indoctrination objection claims that all education involves conveying beliefs which the learner cannot evaluate but must accept. This undermines the ideals of autonomous justification.

Autonomy and rationality are intellectual virtues stemming back to Socrates, corroborated by Kant and the Age of Enlightenment, and held as universal ideals in modern liberal democracies. However, according to Rawls, one of the most influential contemporary liberal political philosophers, making such virtues compulsory may be problematic (Bøyum, 2006). If one assumes that the characteristic of a modern liberal democracy is to acknowledge different worldviews, then one should also consider tolerance towards worldviews that prioritize traditions and authorities like religion over skepticism and critical thinking. Promoting critical thinking as an educational rationale without considering Rawls' argument about tolerance for divergent views is thus naïve according to Bøyum (2006). I will not delve

into that discussion here, except to note that the topic is disputed, and I maintain the view that critical thinking, rationality, skepticism, and intellectual autonomy are commonly accepted as worthwhile educational virtues.

A minimum definition of critical thinking

In spite of the confusion about critical thinking and rationales to support it in education, Ennis (1991/2015) suggests a streamlined conception of critical thinking – "reasonable, reflective thinking focused on deciding what to believe or do," Important here is the basis for decision-making on belief (philosophers' utterance for what to hold as true or right) based on reason.

A committee appointed by the American Philosophical Association suggested a more complex definition: "Critical thinking [is] purposeful, self-regulatory judgment which results in interpretation, analysis, evaluation, and inference as well as explanation of the evidential, conceptual, methodological, criteriological, or contextual considerations upon which that judgment is based" (Facione, 1990, p.3).This definition, as well, highlights evaluating evidence (or reasons) upon which a judgement is based. Further, it includes several aspects that concern both procedure and results. Both definitions echo Dewey, commonly accepted as one of the progenitors of the modern interest for critical thinking, although he used the term *reflective thinking*. Reflective thinking is "active, persistent and careful consideration of any belief or supposed form of knowledge in the light of the grounds that support it and the further conclusions toward which it tends" (Dewey, 1997/1933, p. 9). Dewey's definition, too, highlight reasons – in his words, grounds – to consider a claim tenable.

The approach to critical thinking presented here focuses on the ability to judge a claim's tenability based on reason as the core element of critical thinking. This may include the capacity to distinguish between premises and conclusion in a line of argument, and to evaluate the truthfulness of the premises and the validity of the deduction from the premises to the conclusion. Related capacities include the ability to detect hidden assumptions and unforeseen implications.

Another virtue associated with critical thinking and rational communication is the capacity to express one's line of reasoning in a precise way to avoid ambiguity or misunderstandings. Similarly, the capacity to interpret utterances in a reasonable way is valuable. Lack of precision or reasonable interpretation may lead to disagreement or a consensus that relies on false premises. False disagreement may characterize a situation where two or more discussants seem to disagree, but the disagreement relies more on different

interpretations of a key concept than divergent opinions. False consensus may characterize an analogous situation where discussants achieve consensus by unconsciously interpreting key concepts differently rather than reaching a common understanding.

Rational argumentation

Argumentation is a concept adjacent to critical thinking, emphasized by educators and educational researchers alike (Clark, Sampson, Weinberger, & Erkens, 2007; Erduran, Ozdem, & Park, 2015; Jonassen & Kim, 2010; Leitao, 2000; Noroozi, Weinberger, Biemans, Mulder, & Chizari, 2012; Nussbaum, 2011; Osborne, 2010). Like critical thinking, argumentation is closely related to rationality. Rational argumentation is found to be important to learning for multiple reasons: a) Argumentation helps learners make their knowledge explicit, and thus provides opportunities for explanation and the identification of knowledge gaps; b) Elaborating and adjusting understanding of key concepts may elicit misunderstandings of key concepts in order to promote more developed understanding; and c) Argumentation enables collaboration in learning and developing new knowledge (Andriessen & Baker, 2014).

Since antiquity, Aristotle's thoughts on argumentation have guided Western scholars' ideas on rational thinking. In this tradition, rational thinking consists of drawing valid conclusions from premises. The paradigmatic form of an argument in this tradition is a syllogism, consisting of a major premise, a minor premise, and a conclusion necessarily drawn from the premises:

Major premise:	All humans are mortal.
Minor premise:	Socrates is a human.
Conclusion:	Socrates is mortal.

According to this view, the validity of inference from premises to conclusion is a function of the argument's form. Every deduction, from premises to conclusion, based on a logical form, is valid. This conception of rationality and reason as applications of formal logic has been diagnosed as a mismatch between classical logical theory and what we actually do when we reason and argue rationally (Kvernbekk, 2012). In real life argumentation, we often encounter lines of reasoning that differ from such syllogisms. They may lead us to conclusions that are not *necessarily* but *probably* true or false. Probability may be a function of premises that are more or less conquerable true or inferences from premises to conclusion that are not deductively necessary but rather inductive and contingent. Evaluation of

reasoning will then focus not on deduction from premises to conclusion, but rather evidence and how it backs claims.

Several scholars have challenged the classic conception of logic and rationality. Among the most influential are John Dewey (1920) and Stephen Toulmin, who inspired contemporary argumentation theory (Toulmin, 2003/1958; Toulmin, Rieke, & Janik, 1984). Toulmin's approach to analysis of arguments plays a crucial role in my thesis, and I present his argument model in Study 2 and Study 3.

In line with these ideas, I focus on the justification of claims by providing and evaluating claims' tenability or soundness by scrutinizing reasons (or grounds) as a core dimension in both critical thinking and rational argumentation. This implies that evaluating tenability or "deciding what to believe" is a core element of critical thinking. This is closely related to Johnson and Blair (2006) idea that argument quality is related to an evaluation of the claim's *acceptability*, *relevance* and *sufficiency* (Blair, 2012; Johnson & Blair, 2006).[6]

Additional themes in the critical thinking literature

Even if one accepts that the core element of critical thinking is evaluating claims' tenability (or "assessments of statements," as Ennis put it), major questions about definitions and clarification of the concept remain. To paint a more detailed picture, I will briefly address some additional frequently discussed themes.

Skills vs. attitudes

A central question concerns whether the notion of critical thinking primarily describes a (set of) cognitive skill(s) or if it also contains attitudinal components. Is a person a critical thinker if s/he possess the cognitive abilities to think critically yet rarely uses them? Alternatively, are dispositions featuring a commitment to truth, a willingness to self-correct, or open-mindedness integral and necessary for critical thinking? This was a major question in the American Philosophical Association's attempt to define critical thinking (Facione, 1990), and it has intrigued scholars. Other attitudinal characteristics associated with critical thinking are a willingness to rely on one's own judgments and a critical appreciation of one's intellectual autonomy (Siegel, 1988/2013). This includes a reasonable skepticism towards authorities, which means never to rely on an argument solely based upon the authority behind

[6] This resonates with Næss (1975) instruction to evaluate arguments by assessing *relevance* and *tenability*.

it, except to evaluate its claims' tenability. The commitment to truth and better understanding must be pursued, even when uncomfortable. In some cases, willingness to accept the consequences of one's rational conviction may be linked to critical thinking. Another virtue that has both cognitive and attitudinal components relates to identifying and acknowledging different perspectives or vocabularies. This includes the capacity to reflect on and problematize the contingencies of one's own perspective including its weaknesses and strengths. The use of tendentious and intentionally biased arguments distorts critical thinking, as does rhetorical strategies that intend to lead other discussants astray (such *ad hominem, ad populum,* or straw man arguments). Critical thinking may include both knowledge to identify and expose faulty argumentation and a commitment to desist from using it.

Generic vs. domain-specific

A central controversy of critical thinking research concerns whether critical thinking consists of generic principles that can be learned in one context and transferred to another or whether standards for good thinking are domain-specific and consequently should be learnt integral to a specific area of knowledge. Along with the presentation of what I labelled "a minimum definition of critical thinking" above, proponents of the generic view maintain that the core of thinking critically is to evaluate the tenability of claims. On the other hand, proponents of the domain-specific view (McPeck, 1990; Moore 2004; Moore, 2011a) believe that learning to think critically can only be learnt in relation to a specific subject. From this, they claim a subject-specific approach to critical thinking is more futile. Toulmin (2003/1958)'s argument model (see chapter 1, Figure 3) is an attempt to reconcile the schism between a generic and domain-specific view, as it describes generic aspects of any rational argument while specific *backings* are typically domain-specific.

Thinking as an individual vs. collaborative phenomenon

The cognitive sciences commonly conceptualize thinking as individual cognitive processes. However, in social conceptions (Linell, 2009), thinking designates the act of placing ideas, arguments, or propositions in a *space of reasons* (Sellars, 1997). Such a space is neither individual nor private by nature; rather, the notion highlights that rationality has both social and normative features (Siegel, 2012). Even thinking alone, discussing with oneself is a peculiar variant of the more basic function of discussing with others. Considering reasoning as placing arguments in a shared *space of reasons* deviates from the idea of it as an application of common standards.

The research literature on critical thinking in OEDs I comment on in this thesis sparsely address general research on critical thinking as an educational virtue.

Critical thinking and rational argumentation in my thesis

The aspects of critical thinking and rational argumentation in my thesis emphasize the evaluation of claims' tenability, i.e. how arguments are justified. In Study 1, I suggested the minimum conception of critical thinking is in line with the idea that evaluation of claims' tenability is the core feature. I did this to suggest an alternative to the conceptualization of critical thinking in the Community of Inquiry model. Further, I used this minimum conception to discuss whether the Community of Inquiry model's indicator of critical thinking (progress through phases of an inquiry process) truly represents critical thinking.

In Study 2 and 3, I turned towards Toulmin's argument model. According to Toulmin, a complete argument consists of a claim, grounds that back the claim, warrants backing the grounds, a qualifier that states the relative strength of the argument, and relevant rebuttals to the claim. Central to this model is the idea that an arguments' tenability is a function of how a claim is backed. As such, this model focuses on the core element of critical thinking. In Study 2, I use Toulmin's categories to analyze empirical material from an online philosophy course. My analysis demonstrated that this model does not provide a straightforward analysis of students' critical thinking. In Study 3, I reviewed research literature that employs Toulmin's argument model as a means to analyze students' argumentation in OEDs, and I discuss its affordances and constraints.

2.2. Digitalization of education

The phenomena I investigate in this thesis are research tools for OEDs. It is important to note that my thesis does not intend to develop new knowledge on digital tools and their affordances. My thesis aims to contribute to the quest for better tools for researchers (or at least point to some constraints of established tools) and hopefully enable a better understanding of the affordances of digital communication. In the next few pages, I will outline some issues in the landscape of digitalization of education and corresponding research, and then I will zoom into research on OEDs. In this section, I sketch a funnel where I move from the rather general topic of expectations of digitalization to transform education to the narrower question of how to operationalize critical thinking to analyze OEDs.

Education has greeted digital tools, previously labeled as Information and communications technology (ICT) or simply *computers*, with great expectations. As digital tools and workflows have changed many areas of society during the last 30 years, schools have waited for the marvelous new era to dawn. According to Pedró (2012), "evangelists" and "catastrophists" have dominated debates about the use of digital technology in education which has led to oversimplification as broad general concepts (ICT and digital tools are examples) have dominated the debate.

A major concern both for research and policy is the assumed slow uptake of digital tools in (higher) education referred to as "teachers' non-use of technology" (Selwyn, 2016, pp. 102-108). This has been attributed to their age, lack of competence or interest, or fear of change in status privileges (Selwyn, 2016). One drawback of using broad and generic concepts of ICT is that one fails to observe that some digital tools are ubiquitously accepted (like word processing, Internet searches, and e-mail), while the uptake of others is slower. The use of ICT entails a risk of missing the idea that that different tools have different affordances for transforming education.

The Norwegian Ministry of Education and Research's "strategy for digitalization of the higher education sector" is one contemporary example of both "evangelist" expectations and an illustration of a very broad concept of digitalization:

> Digitalization has changed all sectors of society and is dramatically changing work, everyday life, communication and ways of interaction. Everything ranging from infrastructure and administrative systems; learning material and research data; forms of teaching, learning, and assessment; ways research are conducted; to the very content of the education; and even how higher education institutions interact with society and business life is affected by the possibilities technological developments afford. We live in a time where digitalization and new platforms are having a massive impact on the higher education sector; for both research and education, ICT solutions will contribute substantially in the years to come. Opportunities for reshaped processes of education and research, and new forms of organization and communication will be created through digitalization (Kunnskapsdepartementet, 2017, p. 5, translation my own).[7]

[7] It is not easy to give a modest translation of such a euphoric text. I therefore include the original Norwegian text here:

Noteworthy is that this strategic document only addresses digital tools and digitalization as a generic phenomenon without addressing their myriad functions and significance for the educational sector. Such generic approaches to digitalization fail to describe the various affordances of digital tools.

An alternative scope is to address the different functions technology may play out. In an editorial for American Journal of Distance Education over 30 years ago, Moore (1989) pointed out that imprecise and general concepts hampered communication about, and development of, distance education. According to Moore, interaction is one concept that carries so many different meanings it appears almost useless. To clarify, Moore distinguished three types of interaction afforded by digital technologies:

1. Learner – content interaction
2. Learner – instructor interaction
3. Learner – learner interaction

Digitalization has the potential to transform all these types of interaction. For instance, what counts as learning content and the ways learners interact with content may be transformed. Digital tools enable more ways of interaction between learners and the learning content than reading, listening and writing. More dramatically, when boundless information and supplementary learning material is ubiquitous, learners' interactions may also consist of searching for, critically evaluating, and choosing learning material.

I have briefly depicted this situation to bring to the fore what I believe is missing in the general debate; namely, an understating that digital tools are not one single phenomenon and rather deserve to be discussed with greater specificity. I will now turn to OED as one specific way to use digital tools to facilitate interaction among learners.

Digitalisering har endret alle samfunnssektorer og er i ferd med å endre dramatisk måten vi arbeider, lever, kommuniserer og samhandler på. Alt fra infrastruktur og administrasjon, læringsmateriell og forskningsdata, undervisnings-, lærings- og vurderingsformer i utdanningene og måter forskning foregår på, til innholdet i selve utdanningen, forskningen og hvordan UH-sektoren samhandler med samfunn og næringsliv, blir berørt av de mulighetene de teknologiske endringene skaper. Vi er inne i en tid der digitalisering og nye plattformer med stor kraft er i ferd med å få enda større betydning for sektoren, og både utdanning og forskning er områder der IKT-løsninger vil bidra mye i årene som kommer. Gjennom digitalisering vil det skapes muligheter for nye og endrede lærings- og undervisningsprosesser og nye organisasjons- og kommunikasjonsformer (Kunnskapsdepartementet, 2017).

Digital asynchronous interaction among students

Digitalization affords students the possibility for asynchronous interaction, which is embraced by advocates of constructivist approaches to learning. Digital media enables discussions on a number of platforms such as social media sites, discussion boards, discussion pages in wikis, and comment fields on blogs and webpages. Digital discussion is an important feature of a variety of online platforms and may vary from informal chats in social media to structured exchanges of knowledge and collaborative development of knowledge.

In the early history of online learning and OED, distance education was the primary context. Online discussions were seen as a means to overcome geographical distance and facilitate interaction among learners in different locations. Previously, online conversations have become ubiquitous and is commonly used in both on- and off-campus education. Over the past decades, teaching design for distance education and campus education have converged. Designs developed to overcome distance education challenges are proven to improve campus education as well. Online conversations have become ubiquitous and are commonly used in both on- and off-campus education. In addition, widespread participation in social media has reduced barriers to contribute asynchronously in writing.

Salmon (2004, 2013) has described the educational use of online discussion and conversation in a five-stage model. The three most important steps are socialization, information exchange, and knowledge construction. Constructing and developing knowledge, scrutinizing established conceptions, and contrasting conflicting views requires the ability to reflect, think critically, propose sound arguments, and evaluate arguments' tenability. Critical thinking and rational argumentation are thus both prerequisites for successful discussion and competencies that participation in discussions are intended to develop.

Asynchronous, written, online, discussions have several affordances that may enhance both teaching and learning, and possibly contribute to the transformation of how knowledge is conveyed and developed. Writing at one's own pace provides the opportunity to reflect, collect information, and develop lines of argument. The asynchronous mode enables peers to participate at different times. However, this is established modes of writing and collaboration in traditional, offline media like paper letters etc. Online communication like user groups, message boards, discussion boards and other modes of computer mediated communication added new dimensions to well-established modes of written interaction: distribution of messages became easier, and messages became available to larger audiences (Hillen, 2014; Loncar et al., 2014; Thomas, 2013). Furthermore, while face to face discussions tend to be

dominated by a few active participants, online discussions may facilitate more equal access (Yang, 2008). The easy access – to raise questions or topics for discussion, comment and reply, or just lurk and read – has made online discussions a common media channel for different areas like education, customer support, and all kinds of interest groups. Even if user groups and message boards developed when the Internet came into being, they represent a participatory culture later associated with the emergence of Web 2.0 early in the millennium (McLoughlin & Lee, 2010).

Several perspectives on learning advocate the use of online discussion as a means to stimulate learning. From a social constructivist perspective, the results of learning are the learners' ability to participate in practices with peers. *To appropriate* a concept or knowledge mean to be able to employ it appropriately in social practices with peers. To know something is, according to this perspective, the learner's ability to respond adequately in a discussion or other kind of interaction, rather than a mental state within him or her. Knowledge is thus typically established and co-constructed in collaboration with peers. Online discussions can serve as important arenas for learning. Wise and Paulus (2016) observed that social constructivism is the dominating theoretical perspective in the research field. Even though, other theoretical perspectives are used to shed light over OEDs and learning.

From cognitive perspectives on learning (Kirschner, Sweller, & Clark, 2006; Mayer, 2008), knowledge is considered as changes in an individual's long-term memory. Such changes take place when known information is connected to established cognitive schemes. Discussions may enhance such processes by providing opportunities for reorganizing and consolidating cognitive schemes. Further, retrieval of knowledge from long-term memory, as when using established knowledge in a discussion, is important to consolidate knowledge.

Research methodology on online, educational discussions

From the earliest phase of Internet and computer-mediated communication in education, researchers have tried to figure out how asynchronous discussion might affect learning, and how such discussions are best arranged and facilitated (see for instance Henri, 1992). Does participation in online discussions really contribute to learning, and how is such learning best accounted for? If learning is best supported by mandatory contributions in discussion forums, should there be small or large groups of students? Teachers as active moderators or letting students take the responsibility? What kind of discussion tasks and different kinds of scaffolding (e.g., use of message labels) should be used?

Researchers use several methods to collect and analyze empirical material to understand how OEDs support learning and how they are best facilitated. Number of postings and length of postings may provide knowledge about activity and engagement. Among the methods used in OED research, there is thematic analysis (transcript analysis based on grounded and inductive approaches), social network analysis (how relations between discussants are established and maintained), pre- and post-tests that aim to identify learning gained, and surveys or interviews of students' and/or teachers' perceptions of learning in OEDs. However, to investigate what actually goes on in the discussions and be able to draw inferences about the quality of discussants' participation, learning, and construction of knowledge, researchers need other methods. Discussion of transcripts is a rich material that may provide a window into processes of learning and meaning making. Since the interaction is virtually limited to digital postings, it can be researched and interpreted as pure text without the risk of missing relevant aspects of non-verbal communication. As described in the introduction, the most frequently applied research method in the field is content analysis based on predefined coding schemes. Several reviews address the need for a critical comparison and evaluation of the vast number of these research tools (Clarà & Mauri, 2010; De Wever et al., 2006; Weltzer-Ward, 2011; Wise & Paulus, 2016). My project in this thesis, as stated in the introduction, is to discuss the adequacy of two frameworks for content analysis of OED.

Findings in the research field

Despite the enthusiasm about how OED can stimulate students' knowledge construction and critical thinking, research findings draw a more pessimistic picture. Several articles (Martono & Salam, 2017; Rourke & Kanuka, 2009; Schindler & Burkholder Jr, 2014) concluded that high levels of critical thinking are rarely achieved in OEDs. One question to be raised – that actualizes the discussion of my thesis – is whether such disappointing research findings are a result of weak results or a consequence of deficient research tools (Rourke & Anderson, 2004).

Other areas of research on digital tools in education reveal that the great expectations are hard to substantiate with research. In a meta-study of variables associated with achievement in higher education, Schneider and Preckel (2017) found that use of technology was commonly found to have no (or a small) effect on student learning while variables like social interaction and assessment commonly had a medium to large effect. Låg and Sæle (2019) faced similar results in their meta-study of students' learning from flipped classroom

designs. They expected to find positive effects from flipped classrooms on students' learning and satisfaction, but such effects have been hard to document. This illustrates how positive outcomes related to the great expectation to educational technology are hardly documented by research.

3. Methodology

In this thesis, I aim to evaluate critically the adequacy of two theoretical frameworks – and their coding schemes – used to analyze how OED participation may enhance students' capacity to think critically, articulate, and evaluate rational arguments. As described in the introductory chapter, reviewers (De Wever et al., 2006; Weltzer-Ward, 2011) have raised concerns about validity and remarked that researchers insufficiently build on previous research and compare research findings. What are suitable research methods and design to pursue this concern?

In this chapter, I will first present some general research virtues that motivates my study: generativity (i.e. building on previous research) and validity. I will sketch how these dimensions of quality motivates the studies I have done. Further, I will discuss how my study strives towards generativity and validity. Coherence between research questions, methods, and conclusions is an overall quality of research designs aiding researchers to draw valid inferences. I will present and discuss how I attempt coherence in research design in this thesis, which includes a discussion of the methods I have employed.

Next, I dig into the philosophy of science. A common approach considers ontology and epistemology (metaphysical theories about what reality is and how knowledge is possible) fundamental. Within this approach, the schism between (post-)positivistic approaches and constructivism is central. I present a pragmatist approach to the philosophy of science. Pragmatists reject metaphysical approaches and enable a combination of different methodologies, which is thereby appropriate for my project discussing frameworks' adequacy using arguments based on several approaches.

Finally in this chapter, I discuss ethical considerations I have made through the process of developing this thesis. I structure the section according to considerations towards 1) the research community, 2) the informants, and 3) society in general.

3.1. The ideal of standing on the shoulders of previous researchers

The OED field of research is characterized by a myriad of different approaches, and researchers have put forward several frameworks and coding schemes for analysis of discussion transcripts. De Wever et al. (2006) addressed a concern for weak theoretical backing of frameworks and coding schemes used in the research field. My study corroborates

this concern. Further, De Wever et al. (2006) argued that a lack of solid theoretical basis and validated coding schemes hamper research to move beyond data gathering. In the following paragraph, I will introduce the concept of generativity to describe how research should build on previous research and discuss why this is urgent in this thesis.

Central to scientific activity is to build on previous research, metaphorically to aim to look further by standing on the shoulders of giants.[8] In their seminal paper on PhD education in educational research, Boote and Beile (2005) claimed that:

> To advance our collective understanding, a researcher or scholar needs to understand what has been done before, the strengths and weaknesses of existing studies, and what they may mean. A researcher cannot perform significant research without first understanding the literature in the field. Not understanding the prior research clearly puts a researcher at a disadvantage (Boote & Beile, 2005, p. 3).

A key concept in their argumentation is g*enerativity*,[9] defined as the ability to build on previous research. This implies a call for researchers to locate their work in a theoretical context; identify what is established knowledge within an area, and what still needs to be researched; define the scope of a study according to established knowledge; scrutinize the adequacy of research methods; and, finally, evaluate claims made by previous research.

According to Boote & Beile (2005), educational researchers investigate complex problems often situated in messy contexts. The concepts used in educational research are sometimes vague; divergent theories and explanations are offered, and even sometimes used together. This messy and complex situation make generativity and rigorous literature reviews overly important in educational research.

Reviewing previous literature is the systematic approach for researchers to climb upon the shoulders of giants before us, and thus contribute to a better understanding of what we are studying. Boote and Beile (2005) argue that reviewing literature deserves a more prominent place in doctoral training in educational research. Further, they emphasize that literature reviews should be comprehensive, thorough, and exhaustive. According to them, PhD-students should "thoroughly mine the existing literature" (Boote & Beile, 2005, p. 7).

---

[8] The metaphor is commonly assigned to Isaac Newton, yet its motive has been known since medieval times.

[9] Boote and Beile use the term generativity to designate the ability to build on previous research. However, the term is commonly used to designate a concern to contribute to the next generation. See https://www.merriam-webster.com/medical/generativity (January 17, 2020).

Generativity in my study

One motive for my study is the weak generativity assumed in the research field. With a myriad of frameworks and coding schemes, and scarce discussion or comparison between different approaches, it is challenging to summarize how research has led to a better understanding of how OEDs may improve education. I intend to contribute to generativity in the research field by discussing and criticizing underlying assumptions. How do I demonstrate and ensure generativity in my own research?

In the introduction section of this extended abstract, I sketch out the research field and problems my thesis investigates. I build on previous reviews and overviews to articulate my aims and research questions and to focus on the latter as both filling a gap identified in the research field and problematizing assumptions upon which previous research relied. To establish the research needs for the studies, I explicitly relate my research to previous research and thus aim to contribute to the collaborative endeavor for better understanding. I base my introductory description on secondary sources, previous reviews, and commentary studies. For my purposes, this overview of the research field was more fruitful than exhaustive studies of primary sources would be.

In Study 3, I present an overview of studies that employ Toulmin's argument model to analyze OEDs. According to Grant and Booth (2009), this overview can be characterized as a critical review. Gough, Thomas, and Oliver (2012) recommend alignment in review studies according to aim, research question, literature search, and categorization of included studies. Since my focus was narrow (How is Toulmin's model used in the research field? Are coding schemes based on Toulmin's model adequate as measures for the quality of argumentation?), I omitted the mapping of research questions and empirical findings in my categorization of the studies included. Conducting my literature search, I encountered obstacles identifying relevant studies due to how library databases index educational research. Because of that, I based parts of the literature search on scanning results from a previous review (Noroozi et al., 2012) and hand search. Nevertheless, my literature search mapped the research field for my purpose and provided an overview that nuances claims repeated in previous reviews (Weltzer-Ward, 2011; Wise & Paulus, 2016). A striking observation I had when reviewing the field is that several scholars (Andriessen & Baker, 2014; Noroozi et al., 2012; Nussbaum, 2011) have pointed to weaknesses in using Toulmin's model for assessing argument quality, yet this is scarcely taken up in the research field, corroborating the impression of weak generativity therein.

Boote and Beile (2005) suggested coverage and rigor as criteria for a successful review. I do not meet these criteria completely in my thesis, as I build on secondary sources, such as previous reviews. Nevertheless, in contrast to comprehensiveness, thoroughness and exhaustiveness, Maxwell (2006) suggested *relevance* as the key criteria for literature review. He distinguishes between reviews *of* research, and reviews *for* research. According to Maxwell, a comprehensive inventory of research publications in a defined research field is neither necessarily interesting nor feasible. What is needed is an overview that locates gaps or problems. One danger of aiming for comprehensiveness by narrowing scope is a risk of becoming "a prisoner of the theoretical or methodological perspective that dominates this literature, and fail[ing] to see alternative ways of conceptualizing or studying the issue or problem" (Maxwell, 2006, p. 29). The review part of my thesis does not provide exhaustive overviews of the research field as suggested by Boot and Beile, yet my overviews enable me to identify and discuss urgent questions from previous research, thereby contributing to generativity by criticizing specific underlying assumptions.

3.2. Validity

In my thesis, the concept of validity plays a crucial role. The motive for my study is the concern that frameworks used to analyze the quality of students' contributions in OEDs do not adequately measure (thereby enabling inferences about) what they intend to measure. Validity is a key measure for research quality and describes whether the inferences that researchers draw are credible and trustworthy. In empirical research, validity concerns the quality of the inferences that researcher can draw from data (AERA et al., 2014; Kleven, 2008; Kleven & Hjardemaal, 2018; Maxwell, 2013; Messick, 1995; Shadish et al., 2002). In other words, questions about validity concern how researchers might draw fallible inferences, and if alternative inferences might be interesting and trustworthy (see for instance Kleven & Hjardemaal, 2018; Maxwell, 2013, pp. 121-123).

In argumentation theory, Blair and Johnsen (Blair, 2012; Johnson & Blair, 2006) have suggested *relevance*, *acceptability* and *sufficiency* as criteria to evaluate whether a claim is well supported and thus tenable. To consider the validity of a claim made by research is to consider whether it is relevantly, acceptably and sufficiently backed up.

In the following section, I will first sketch conceptions of validity relevant to my critical evaluation of frameworks used to analyze critical thinking and rational argumentation.

In the next section about coherence and research designs, I present how I have designed my research to ensure validity for the inferences I draw.

Shadish et al. (2002) presented an account of validity that has become influential for quantitative research in psychology and educational studies. It is suitable for the frameworks I scrutinize, since they are commonly used within quantitative educational research. Kleven (2008); Kleven and Hjardemaal (2018) have suggested that this conception of validity is relevant to qualitative research as well: Shadish et al. (2002)'s validity dimensions can fruitfully be adjusted to discuss quality of research based on several research approaches.

Shadish et al. (2002) claim that the inferences a researcher draws can be evaluated according to four dimensions of validity:

1. Statistical validity concerns whether co-variation between variables is statistically significant, or merely random.
2. Internal validity concerns whether and how co-variation between variables can be inferred to indicate a causal relationship between the variables or rather represent spurious or accidental co-variation.
3. Construct validity concerns how observable indicators represent the construct they are intended to represent. This is the particular aspect of validity I focus in my thesis, and I will return to this below.
4. External validity concerns whether inferences drawn from a study are relevant for other contexts. In other words, to which extent are findings and inferences generalizable?

These dimensions form the basis for a set of questions that can be raised towards research claims based on both quantitative and qualitative methodologies (Kleven & Hjardemaal, 2018):

1. Do research findings represent an important relation, or do they merely express a random coincidence?
2. Do correlations represent a causal (or at least interesting and relevant) relationship? More broadly, which alternative explanations are possible and trustworthy?
3. Do observations really describe or measure what they claim? Are the observed indicators and characteristics representative of and relevant for what they are taken to represent?
4. In which contexts are the findings relevant? Are findings generalizable to other contexts?

Maxwell (2013, pp. 124-125) adds two threats to validity in qualitative research. They may also influence on other research approaches as well:

5. Reactivity: How have informants been influenced by the data collection?
6. Bias: How have the researchers' bias influenced data collection and analysis?

The relevance of each issue may vary from one research approach to another. For instance, the question of generalizability (4) may be less relevant for idiographic research aiming at in-depth understanding of a single case. Quantitative research typically aims for generalizability by analyzing data that are claimed to be representative. In qualitative research, the question of generalizability is sometimes transferred to the reader, who is handed the challenge of considering whether findings are relevant outside their original context (Kvale & Brinkmann, 2009). Nevertheless, considering which contexts the findings are represent and if there are any insights that may reach beyond that is instructive for research from several perspectives. Considering internal validity understood as causal mechanisms (2) has limited relevance for interpretative research based on constructivist approaches. However, pondering alternative interpretations and explanations may be worthwhile regardless of research approach.

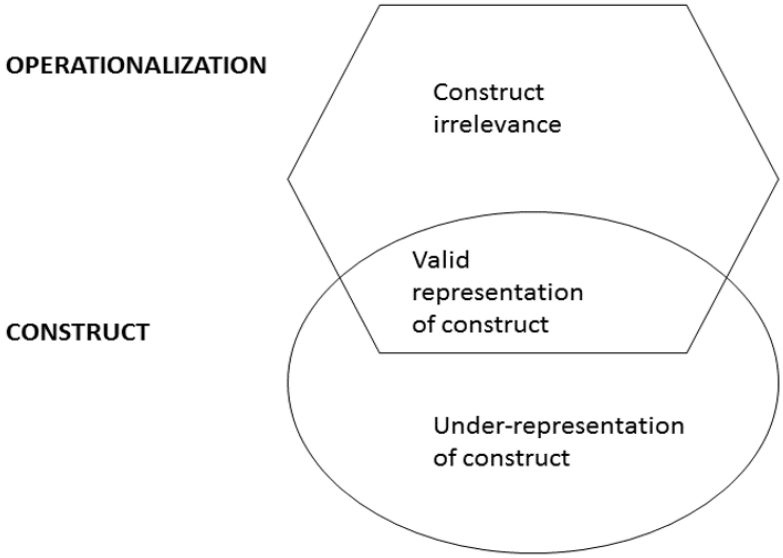
Construct validity and frameworks for analyzing OED

According to Shadish et al. (2002)'s validity dimensions presented above, the questions of my thesis concern construct validity. When researching abstract phenomena that cannot be directly observed, such as rational argumentation and critical thinking, researchers need to operationalize these constructs into observable indicators. Such indicators may be more or less representative of – and more or less relevant for – the construct they are intended to describe (AERA et al., 2014; Kleven, 2008; Messick, 1995). *Construct representativeness* and *construct relevance* are two aspects that, together, comprise construct validity (Figure 5).

Construct representativeness describes whether important features of the construct are represented by the indicators. When an operationalization fails to identify necessary, sufficient or at least central features of the construct as indicators, the operationalization is considered construct under-representative (i.e. it has *construct deficiency*). If necessary, sufficient or necessary features are missing among the indicators, there is a risk for false negatives, that is, to falsely code something as not representing the construct even if does represent the construct by any reasonable definition. For instance, if use of Latin terms is held as the sole indicator of critical thinking, then will an argument that questions justification (a commonly held important feature of critical thinking), yet lacking Latin terms, be coded as not representing critical thinking.

Figure 5

Construct validity. Based on (Kleven, 2008)



On the other hand, construct relevance describes whether the indicator (or set of indicators, i.e. the operationalization) identify relevant, that is, necessary, sufficient, or at least central features of the construct. Coding based on irrelevant indicator(s) may lead to false positive results, that is, when something is falsely coded to represent the actual construct. Irrelevant indicators entail *construct contamination*. For instance, if use of Latin terms is (irrelevantly and falsely) taken as indicator of high quality of arguments, then coding based on this indicator will result in that weak arguments (by any other reasonable definitions of argument quality) containing Latin terms will be falsely coded as representing high quality.

In the examples above, the use of Latin terms is held to be an indicator (yet irrelevant) of rational argumentation and critical thinking, while asking for justification is held to represent a central feature. The crux of developing a coding scheme is to identify indicators that are both representative for central features of the construct under investigation, and observable when coding.

In my thesis, I ask whether progress through phases of an inquiry process is an adequate indicator for Cognitive Presence (i.e. critical thinking) and whether the presence of Toulmin's argument components is an adequate indicator of rational argumentation. To anticipate my conclusion, I find both these indicators to represent weakly the constructs they are intended to represent. Further, I find these indicators to be not necessarily relevant indicators of the constructs they are claimed tor represent. Being *not necessarily relevant* does

not mean that such features are consequently irrelevant – in many cases, these indicators may describe worthwhile features. My argument is, however, that these features do not serve as necessary or sufficient indicators of critical thinking and argument quality, and thus poorly discriminate between strong and weak arguments.

Different conceptions of construct validity

In my Study 1, I used Shadish et al.'s (2002) conception of construct validity and Messick (1995) concepts of construct relevance and construct representativeness for my analysis. I found the concepts of construct relevance and construct representativeness useful categories to discuss how far the operational definition (progress through phases of an inquiry process) works to indicate the Cognitive Presence construct and thus critical thinking.

Nevertheless, some research methodologists also use the term construct validity with a different meaning. Instead of describing, as above, a match between the conceptual/ theoretical definition and its operational definitions, construct validity is used by some research methodologists to note the extent to which a measure is related to other measures (constructs) in a way consistent with hypotheses derived from theory (Neuendorf, 2002, p. 117). According to Creswell (2014a, pp. 181-182) construct validity can be determined through statistical analysis.[10] This use of the concept construct validity differs from Shadish et al. (2002)'s and from my own in the first study.

After my first study about construct validity in the Cognitive Presence construct was published (Breivik, 2016), Caskurlu (2018) published an article titled "Confirming the sub-dimensions of teaching, social, and cognitive presences: A construct validity study." This article used statistical analysis of the interrelations between Cognitive Presence and other dimensions of the Community of Inquiry framework to demonstrate construct validity. Yet there was no discussion about how far the indicators for Cognitive Presence described what they were intended to do nor was there a discussion of indicators for other constructs present in the Community of Inquiry model.

Due to the various meanings associated with the term, I decided not to use *construct validity* in the subsequent studies for this thesis. I did this to avoid going into discussions about several conceptions of validity, and to thus focus on my main concern: Does the Cognitive Presence construct from the Community of Inquiry framework and frameworks

[10] "Through statistical analysis you can: 1) see if scores to items are related in a way that is expected, 2) test a theory and see if the scores, as expected support the theory […This] relates to the traditional notion of construct validity" (Creswell, 2014a, pp. 181-182).

building on Toulmin's model analyze critical thinking and rational argumentation as they were intended? I decided to use the less precise terms of *adequate*/ *adequacy* rather than the more precise yet polysemantic term of *construct validity* when revising my research questions. By asking about adequacy, I attempt to question the suitability (relevance and representativeness) of the indicators from the Community of Inquiry's construct Cognitive Presence and the categories from Toulmin's model for analyzing critical thinking and rational argumentation.

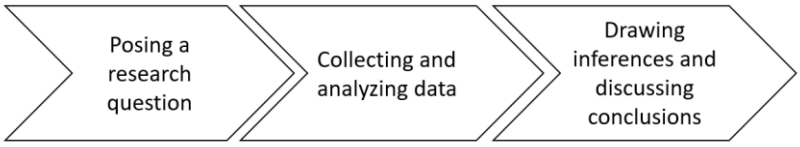
3.3. Research design and coherence

According to Creswell (2014a, p. 17) and most research methodologists, the basic stages in a research process are:

1. Posing a research question.
2. Collecting and analyzing data.
3. Drawing inferences and discussing conclusions.

Figure 6

Coherence in research design.



Validity concerns the quality of the inferences drawn by research. Focusing on coherence between research questions, methods for collecting and analyzing data, and inferences (Figure 6) is one approach to ensure validity. If there is incoherence between the claims made, the ways these claims are backed, there is a risk that the claims have backing that is *irrelevant*, *unacceptable* or *insufficient* (Blair, 2012; Johnson & Blair, 2006). Examples of such a lack of coherence may be attempts to answer conceptual questions using empirical methods, or vice versa (e.g., answering the empirical question "How to best enhance learning?" by conceptually analyzing enhancement strategies).

In the coherence model, the aim and research question are pivotal. In an ideal research process, the aim shapes the literature review, which forms the basis for research questions, methods of data collection and analysis, and thus enables findings – providing a basis for conclusions that can be drawn. However, the aim and research question may stem from either

queries in a research field, challenges in a practical enterprise, or both. As described in the introduction chapter, the research questions in this thesis, even if theoretical, emerged out of a practical situation: I was to conduct analysis of empirical material from online discussions. For that purpose, I needed an appropriate framework for my analysis. Reviewing the research literature, I recognized an already identified gap in the research field – the need to scrutinize acknowledged research instruments. Further, I suspected that these research instruments, even if well-established and acknowledged in the field, relied on erroneous assumptions.

The coherence model can hardly be claimed to be descriptive; the research questions are commonly written or adjusted late in the process, especially in qualitative research. Maxwell (2013) explicitly warns against adhering too strictly to a research project's initial questions and design or following plans too rigorously. This may lead to a certain blindness towards knowledge and ideas generated during the process and increases the risk of asking the wrong questions. Thus, the coherence model does not imply that the phases of the research process should be performed in a timely order, yet coherence between the research aims and the research question(s), methods, findings, and conclusions parallels the coherence between a conclusion and the premises that support it. In some research paradigms, like grounded theory and some inductive strategies, an unbiased approach to empirical reality is the research ideal, although there needs to be coherence between conclusions and what serves as their backing.

Pointing to the value of coherence – as is the message in Figure 6 – may sound trivial. Nevertheless, think of the idiom about how problems appear to you when your only apparent tool is a hammer. Similarly, think of the joke about the man searching for his lost keys in the light under the streetlamp – because that's the most convenient place to search – not the place where he lost his keys. These parables (both inspired by Maxwell, 2013) remind us how researchers may prefer handy research methods out of convenience or habit, not alignment to research aims and questions.[11] Sometimes the temptations arise among researchers to draw conclusions that transcend what they have empirical backing for (among other kinds).

According to some research methodologists, (Krumsvik, 2016; Maxwell, 2013) scientific theories can be divided into variance theories and process theories. Variance theories explain relations between dependent and independent variables, aiming to identify causes and effects. Typically, research questions focus on co-variation, causality, or quantification: "What is the relation between x and y?" or "Does x lead to y?" or "To what

[11] The relatively sparse use of observational studies in educational research may result from this tendency to use the most convenient research methods (Wells, Kolek, Williams, & Saunders, 2015)

extent is x representative for y?" Consequently, quantitative methods are most appropriate. Process theories, on the other hand, focus on understanding contexts, meanings, motives, and experiences: "What does x mean (for y)?" or "Why does x want y?" or "How does x experience y?" As such, qualitative approaches are appropriate for data collection and analysis. Ensuring coherence between type of question, theoretical underpinning (process vs. variance theory), and methods (qualitative vs. quantitative) is vital. These elements are commonly related to specific philosophy of science positions, which I will return to in a later section. Before that, I will comment upon the kinds of questions I raise in this thesis, and which methods are suitable to ensure coherence between the questions I raise and how I intend to answer them.

Coherence in my study

The aim of this thesis is to discuss the adequacy of two theoretical frameworks and their coding schemes to understand how participation in OED enhances students' capacity to think critically and articulate and evaluate rational arguments. To discuss this, I raise the following questions:

1. How is critical thinking analyzed in the Community of Inquiry framework and how adequate is this operationalization?
2. How do frameworks building on Toulmin's argument model analyze quality of argumentation, and how adequate is this way of operationalizing rational argumentation?

What kinds of research design and methods are appropriate to answer such questions? The questions I raise are neither variance questions (about relations between measured variables) nor process questions (about actors' experience or motives, or specific contexts). Rather, they inquire about the relevance and representativeness of certain indicators to given constructs. This call for a conceptual and philosophical analysis.

Philosophy of education

Philosophy of education is not one unified enterprise. Frankena (1968) established a distinction between analytic and normative philosophies of education:

> The philosophy of education […] may be either *analytical* or *normative*. It is *normative* insofar it is concerned to propose ends or values for education to promote, principles for it to follow, excellences for it to foster, methods, contents, programs, etc., for it to adopt or employ, in general or in specific situations. It is *analytical* as it is concerned merely to analyze, clarify, or elucidate, or to criticize and evaluate our thinking about education

– the concepts or terms we employ, the arguments we use, the assumptions we make, the slogans we proclaim, the theories we formulate" (Frankena, 1965, p. 8).

The motive for analytic philosophy of education is thus to identify and clarify confused or unclear notions or hidden assumptions. On the other hand, normative pedagogical philosophy aims to propose values and ideals for educational practice. From an analytical perspective, philosophy can be understood as discipline without its own domain, yet it is an intellectual activity that identifies (more or less hidden) assumptions within other disciplines and analyzes and clarifies concepts. Understood this way, philosophy is not queen of all sciences but rather dependent on symbiosis with other disciplines. Empirical researchers may stumble upon problems defining the concepts they use. What is justice in education? What is (how to define or operationalize) critical thinking? Even more seriously, researchers may face problems identifying and delineating their object of study or scientific discipline: What is pedagogy, and should it be defined by the object of study or its methods? What is learning? When empirical educational researchers face questions concerning object identification, this calls for a philosophical approach. Such questions cannot be answered by empirical analysis alone since important questions concern what qualifies as empirical data. Facing questions that cannot be solved by established empirical approaches, philosophical reflection and conceptual analysis might be fertile. One possible criterion of success for philosophical approaches applied to the concepts from empirical research is that this approach provides arguments or clarifications that are found relevant within the discipline where the questions occurred.

The puzzlement on how to best analyze OEDs and concerns about the approaches used in the research field calls for conceptual and philosophical analysis of key concepts. Thus, my motive in this thesis is primarily analytical, to clarify and evaluate, rather than promoting a certain approach to education. This influenced my choice of research methods. Despite my analytical motive, I will return to implications of how to enhance competencies in critical thinking and rational argumentation in the final chapter.

3.4. Methodological toolbox

Research methods describe the procedures that ensure the search for knowledge is rigorous, thus supporting the validity of the inferences drawn. How have I pursued the research questions in this thesis? In the following section, I will present some strategies and tools I employed in my analysis and discussions: identifying arguments and condensing key

ideas, tracing philosophical ideas, and testing coding schemes on empirical material and thought experiments.

Identifying and scrutinizing arguments

One important mode of philosophical work is to analyze arguments. Presenting clear claims (or conclusions) and their backings (premises for conclusions) make arguments transparent. Transparency enables inquirers to scrutinize concepts, assumptions and backings, thus making it possible to assess whether the backings are acceptable, relevant, and sufficient (Johnson & Blair, 2006). Dewey expressed an analogous idea, quoted previously in this thesis (section 2)) when he said reflective thinking is "active, persistent and careful consideration of any belief or supposed form of knowledge in the light of the grounds that support it and the further conclusions toward which it tends" (Dewey, 1997).

Dewey's words focus on both supporting grounds and further implications. When I discuss how the frameworks I evaluate operationalize critical thinking, I consider both the professed reasons and possible implications of analyzing OEDs by use of the actual framework. In addition to arguments and lines of reasoning themselves, the terms we use also need to be scrutinized to enhance transparency, which is a motive in this thesis. Key terms like *critical thinking* and *construct validity* are commonly used in an ambiguous way. Terms like *ICT* and *digital tools* (see also chapter 2) are used with different meanings, sometimes vague. In the Norwegian context, Næss (1975, 2005)'s classical book on argumentation for the introductory university course in philosophy highlighted the benefits of analyzing concepts and assessing arguments. I aim to display ambiguities in the use of terms and explain how I employ terms accordingly.

Condensation and straightforward formulation of key ideas

Analyzing a concept or considering a claim requires a definition of the concept or formulation of the claim that is condensed, transparent and unambiguous. Such condensed to-the-point formulations need to be un-controversial and accepted by the proponents to be viable for further analysis and discussion. (Næss, 1975, 2005). Condensing and summarizing key ideas that is inconsistent with the views of those who held them risks falling into the fallacy of strawman argumentation.

When Garrison et al. (1999); Garrison (2011); Garrison et al. (2001) and other proponents of the Community of Inquiry framework present the Cognitive Presence construct and related indicators, they refer to the so-called Practical Inquiry model presented in chapter

1. They claim that the model integrates elements of individual reflection, collective discourse action /practice and perception of ideas. The model is rather complex and it does not illustrate Dewey (1933/1986) thinking in a straightforward manner. Proponents have broken down this model into categories (triggering event, exploration, integration, resolution) with associated indicators. In Study 1, I summarized this to the claim that the indicator of Cognitive Presence and thus critical thinking is progress through phases of an inquiry progress.

Is progress through phases of an inquiry process a precise and uncontroversial condensation of the Cognitive Presence construct, as I have claimed? My point is backed by a quote from Shea (2010, p. 15): "Deep and meaningful learning does not occur until students move to integration and resolution stages" (p. 15). Garrison, one of the founders of the Community of Inquiry model (Garrison 2017) has commented on my study 1 (Breivik 2016). He did not reject that progress through phases is an observable indicator of Cognitive Presence. (Nevertheless, he rejected my claim that the Cognitive Presence construct is intended to represent critical thinking. I will comment on this later in chapter 5.)

Most researchers who use Toulmin's model for analysis of argumentation in OED maintain that they analyzed argumentation based on Toulmin's analysis of argumentation. Since Toulmin is widely acknowledged, this seems convincing. However, those who perform analysis based on this model do not use the condensed formulation "presence of argument components is an indicator for quality of argumentation". Coding schemes that code for appearance of argument components and sum up an overall score may be reasonably interpreted this way. This is in line with Nussbaum (2011) who sums up this approach as "counting argument components" (p. 100). As I discuss in Study 3 (Breivik, 2020b), counting the presence of argument components is a problematic measure for argument quality.

Tracing ideas back to their origins

When educational researchers work on topics like critical thinking, rationality, and argumentation, the philosophical tradition is a nearby discipline for inspiration. Both the Community of Inquiry framework and frameworks based on Toulmin's argument model borrow authority by referring to modern classics from the discipline of philosophy. Are the uses of these canonical ideas compliant with the originators' motives?

Dewey's (1933/1986) thoughts on education and learning, especially in the book *How We Think* serve as a platform for Garrison and colleagues' Community of Inquiry framework and the Cognitive Presence construct. May a recall of Dewey's project shed light on the soundness of the framework? In Study 1, I demonstrated how Dewey' ideas about thinking

phases reflects his ideas on inquiry as embedded in social and practical contexts. According to Dewey, progressing through such phases does not serve as an indicator for successful thinking.

Toulmin's argument model and his work on argumentation have received wide application, including contexts outside Toulmin's original aims (see for instance Bostrom, 2003; Hegelund & Kock, 1999, see also Toulmin's own remarks in the introduction to Tolulmin, 1958/2003). In Study 3, I demonstrate that Toulmin's intention when suggesting his model was far from counting presence of argument components as indicator for argument quality.

Alignment with the originators' intentions is not necessarily a prerequisite for sound use of ideas. For instance, Toulmin's model is useful in communication studies, far from his original motive. The research field of EdTech is diverse, but comprehension of philosophical theories may sometimes seem dubious. When models from philosophy are validated by appeals to their status in their mother discipline, it may be relevant to trace the origins of such models. In Study 1 and 3, I sketch Dewey's and Toulmin's ideas as a background to how their ideas are used in research on OEDs. Tracing the origin of key ideas – like Dewey's inquiry processes and Toulmin's argument model – does not provide conclusive arguments about the adequacy of certain approaches to operationalize critical thinking and rational argumentation. However, knowing the background of such ideas adds nuance to understanding how their prestige in the discipline of philosophy validates their use in OED research.

Testing coding schemes on examples

One approach to test how a framework and coding scheme serves as a means to analyze discussion transcripts is to apply it on empirical or fictive examples. Clark et al. (2007)' used this strategy when reviewing several frameworks suggested for analyses of argumentation in online discussions. They coded a short excerpt, demonstrating how different frameworks have different affordances.

In Study 2, my strategy is to test a coding procedure on empirical material using categories from Toulmin's model. The aim is twofold: 1) to explore what goes on in the discussions using these categories, and 2) to discuss how such categories work for such analysis. The result showed that use of Toulmin's categories enable analysis congruent to students' and teachers' own evaluations of argumentation in the discussions, but with greater subtlety. I will present my procedures for the empirical analysis in the next section. In Study 3, I use the same method on small excerpts (derived from Internet pages about flat Earth

theory). Here, I use examples to demonstrate how coding based on Toulmin's categories may lead one astray. I comment on this in the subsequent section on thought experiments and counterexamples.

Empirical analysis in Study 2

In Study 2, I employ empirical analysis in addition to the predominantly argumentative and conceptual analysis in the rest of the thesis. I will here briefly describe the selection of data. The steps of the analysis are described in the study.

The aim of the study is twofold: 1) to explore what goes on in OEDs, analyzed using categories from Toulmin's model; and 2) to discuss how these categories work for such analysis. To do this, I needed a suitable data material. At the time, the Faculty for Humanities, Social Sciences and Education at UiT – The Arctic University of Norway promoted flexible introductory courses, which included OED as a learning activity. As an educational consultant (part of the teaching obligation included in my PhD scholarship), I had access to discussion boards from courses in several subjects. OEDs from the introductory philosophy course were suitable for my purposes: knowledge of argumentation is integral to the subject of philosophy, and competency in argumentation was an explicit learning objective for the course. In addition, discussions from this course were rich and lively. The teachers emphasized OEDs in the course design and developed teaching strategies to enhance activity in the discussion boards.

The selection of the data material was based on principles of purposeful sampling (Creswell, 2014a, pp. 228-231), to explore how students presented arguments in good-functioning OEDs. My aim was not to gain generalizable knowledge about argumentation patterns in OEDs, but to explore possibilities related to OEDs as a learning activity. Furthermore, I sought to discuss how Toulmin's categories served as analytical tools. For this, I needed suitable material from lively and rich discussions. After I chose to use this material, I reported the data collection to the Norwegian Data Protection Services (NSD) and acquired students' consent to use their already completed writings for my research.

Thought experiments and counterexamples as argumentative strategy

Thought experiments are a way of exploring possible consequences of principles or a set of categories by imagining how they will behave in a real world application (Maxwell, 2013, pp. 68-72). A counterexample is a kind of thought experiment common in philosophy. Among the most famous is Gettier (1963)'s, used to prove that the famous and acknowledged

definition of truth as "justified, true belief" is insufficient. Gettier examples set up fictitious situation where a person believes to possess justified, true knowledge. In such examples, it is obvious that even if all the three requirements in the definition are fulfilled (i.e. beliefs prove to be both true and justified), they still cannot count as knowledge in any reasonable sense of the term.[12] The imagined examples demonstrate that the definition is insufficient. The mode of counterexamples checks whether a definition enables false positives or false negatives.

In Study 3 (Breivik, 2020b), I test Toulmin-based coding schemes on examples derived from webpages about flat Earth theory with the intent to demonstrate that while the coding schemes may seem sound, they can produce unsound analysis. Even if indicators of hjgh-quality argumentation according to the coding schemes are present in an example, the example can hardly be held to be a high-quality argument. Ergo – the indicators are dubious. Rather than characterizing this demonstration of coding based on Toulmin's categories testing on (quasi-)empirical examples (see above), I characterize it as a thought experiment and a counterexample due to the selection of examples. I selected the examples by expediency to demonstrate a certain point, not to perform a candid analysis of empirical excerpts. Nevertheless, my demonstration shows how coding may led astray.

Summing up

The ways scholars approach conceptual and argumentative analysis are sometimes based on tacit knowledge. In the previous sections, I have explicitly described how I have approached my research questions by identifying arguments and condensing key ideas, tracing philosophical ideas, and testing coding schemes on empirical material as well as thought experiments. I have done this to make my arguments more transparent for review and discussion. Further, the approach I suggest, combining and explicating these approaches may add to the development of methodology for conceptual research.

[12] Imagine you are looking into the woods. You see something moving which you take to be an animal, and you infer that there is an animal in the forest. What you saw was just the wind moving some leaves. In fact, there is an animal, which you did not see in the forest. Do you have knowledge that there is an animal in the forest? Your belief is true (in fact, there is an animal there) and justified (you inferred it from your senses). However, it is merely luck that your inference is true since you inferred it falsely from poor evidence.

3.5. Philosophy of social science

I previously described coherence in research design as an alignment between research questions, methods, and material to answer the questions; and the findings and conclusions researchers can draw. In research design, coherence may also concern philosophy of science assumptions, methodology, questions asked, theoretical underpinnings, as well as findings and conclusions. Philosophy of science assumptions influence the practice of research, even when implicit and hidden (see for instance Creswell, 2014b).

What assumptions about philosophy of science do I pursue in this thesis, and how do such assumptions influence my research? Before I open these questions, I will draw in broad strokes how philosophy of science textbooks for educational research commonly introduce the topic (Crotty, 1998; Krumsvik, 2016; Twining, Heller, Nussbaum, & Tsai, 2017). Assumptions from the philosophy of science may deal with ontology – questions about the nature of reality; epistemology – questions about how knowledge about reality is possible; and methodology – relations between epistemology and procedures to gain backing for knowledge claims. Philosophy of science is commonly considered fundamental in the way that research methods are based on methodology, which is based on a theoretical perspective; which is based on epistemology; which finally is based on ontology. According to this view, a profound conception of the nature of reality's existence is necessary as a foundational base for research. I will comment on this idea of foundationalism below, after presenting very broadly two typical conceptions of classical approaches to philosophy of science (see Table 1): (post-) positivism[13] and constructivism. Both take ontology as a starting point. What kind of knowledge research can aim for, and how such knowledge is possible, is a necessarily a result of one's vision of reality.

[13] The logical positivists aimed to find a demarcation criterion that could distinguish between real science and pseudo-science. They claimed it was verification; namely, a scientific hypothesis should be able to be proven true by empirical evidence. Further, they claimed that observation and data collection should aim to be objective and theory-free. Popper and others have problematized these ideals, noting that some theories have a tendency to self-validate and that the potential for falsification is a better criterion. Further observation and data collection without any direction is a naïve and meaningless idea. Despite the differences between positivists and post-positivists, both hold the idea that science entails objective descriptions and causal explanations of an external and independent reality.

(Post-)positivism

(Post-)positivist ontology claims that that the world exists as material reality independently from us humans as observers. The related epistemology holds that knowledge is true descriptions – like mental pictures or verbal descriptions – that correspond with an

Table 1

Philosophy of science-positions. Based on Crotty (1998), (Krumsvik, 2016), and Twining et al. (2017).

	(Post-)positivism	**Constructivism**
Ontology	One objective reality exists independently from human observers. Reality is characterized by cause and effect relations.	Multiple realities exist. Meaning is essentially a part of reality, and relative to humans.
Epistemology	Knowledge is true descriptions of a human-independent reality.	Interpretation of meaning demands pre-understanding and is thus relative to the knower.
Methodology Based on ontological and epistemological foundations)	Quantitative methods are preferred since these are the least influenced by the researchers' bias. Sharp distinction between context of discovery and context of justification. Only context of justification is scientifically relevant.	Qualitative methods are best suited to grasp meaning and actors perceptions. Inductive approaches enable interpretations based on the actual case, not pre-established theories.
Types of questions	Variance questions: - quantitative relations between independent variables and dependent variables.	Process questions: - qualitative accounts of actors' perceptions and motives within an authentic context.
Types of explanation	Causal explanations.	Interpretations.
Scientific ideals	Value-free and objective descriptions of causal relations to predict future events.	Thick descriptions: Understanding actors' perceptions and motives within an authentic context.

objectively existing reality. The aim of research is to establish objective descriptions of this external reality, preferably descriptions that enable causal explanations and, in turn, precise

prognosis of future events. The preferred methodology is quantitative, as this is considered to be least influenced by the observers' subjective perceptions. A common research strategy is testing of hypothesis, and the distinction between contexts of discovery (creation of a hypothesis) and contexts of justification (testing the hypothesis) is vital. Only the context of justification is relevant for the truth of scientific knowledge. Reductivism, to treat all objects, including humans, as things, and all scientific knowledge as testable generalizations is one variation of this idea. This view implies that social sciences should be conducted in the same way as natural sciences.

The (post-)positivist view implies some troublesome philosophical questions: How is it possible to account for correspondence between essential different entities, i.e. the external world vs. a subject's inner picture of this world? The only way a knowledge-seeking subject can check if her inner picture based on sense perceptions corresponds with reality is to compare it with another sense perception-based inner picture of reality. According to critics, the knowledge humans rely on when understanding our surroundings is more than observations of physical entities. Interpretation of meaning is crucial part of human knowledge which can hardly be accounted for in a positivist approach. Further, maintaining that knowledge consist of inner pictures of reality that correspond with the external, mind-independent reality, within a subject observing her surroundings disengaged and objectively, dismisses that humans are engaged actors, not passive contemplators of their surroundings.

Constructivism

Constructivists, on the other hand, insist that (at least parts of) our world is laden with meaning. The objects that surround us are not merely physical entities like sheets of paper or pieces of metal, they are coins and bills, or jewelry and letters soaked with historical, cultural, and social meaning. Reducing meaning to observable entities, as positivist science suggests, diminishes the option to grasp anything interesting about social phenomena. Meaning cannot be neutrally observed but must be interpreted and understood. This presupposes that understanding is enabled by previous understanding – thus abandoning any objective or neutral approach – and demonstrates how the world around us is not a primarily a world of physical entities but a world of meaning. Since meaning shapes the world and how we perceive it, understanding actors' meanings is more fundamental than explaining causality. The methodology of researching actors' worlds – laden with meaning – is better done by interpreting meaning than objective observation.

For constructivists, it is problematic to hold knowledge as representations of a reality independent from humans. If all (or most of) what is worth researching is soaked with meaning, and meaning demands interpretation (which again is enabled by individual prejudice and historical/contextual horizons), then there is no option to gain objectively true knowledge. Literally, there is no mind-independent reality worth investigating. Further, the methods employed (small-scale interpretation or in-depth studies) have poor options for generalizable knowledge. Some critics claim that this leads to anti-realism (denying that an objective, mind-independent reality exists) and thus relativism (abandoning the possibility of objective knowledge, claiming that all knowledge is relative to personal prejudices).

Combinations of these two approaches are suggested. Typical is the division of natural sciences on the one side (building on some variant of positivism) and the humanities on the other side (building on a constructivist approach). Some scholars aim to include both approaches by holding two different ontologies, claiming that the (post-)positivistic view suits natural phenomena (object ontology), while phenomena laden with meaning (subject ontology) are best researched by interpretative methods (Krumsvik, 2016, pp. 100-101 refering to Searle, 2015). One challenge for such an approach is that phenomena that are obviously based on subjective meaning (e.g. individual political preferences) and thus "ontologically subjective" (Searle's terminology) are on an aggregated level (e.g. electoral behavior) fruitfully studied by methods associated with objectivist ontology. It remains to be explained how a phenomenon changes ontologically from the individual level to aggregated levels. In line with this, critical realists like Bhaskar (2008) claim that to avoid the relativism of constructivism, the discipline of ontology is essential as a foundation for the methodology of social sciences. Bhaskar claims that natural and social sciences share some features, yet different research disciplines study objects of different ontological status.

The scheme presented in Table 1 above is useful to broadly describe traditions in social sciences and educational research. The two traditions hold contradictory, possibly incompatible claims on key issues. Despite this, the two traditions can be lumped together in what Morgan (2007) labelled as the "metaphysical paradigm." Central is the view that methodology is reliant on epistemology and epistemology is reliant on ontology. Ontology is thus held as fundamental.

Where in this landscape of philosophy of science have I situated my thesis, and how do I navigate it? I have previously described my research questions as conceptual or philosophical questions rather than variance or process questions. In the introduction, I stated

my aim to participate in a methodological debate. The research field I comment upon – content analysis of OED – makes extensive use of quantitative approaches and statistical analysis of a phenomenon that concerns individuals' perceptions on a basic level – how they think and present arguments. The approach I employ when analyzing empirical material in Study 2 is qualitative rather than quantitative. The backings I use to support my conclusions are built on conceptual analysis rather than quantitative empirical data.

So, what are the underlying philosophical ideas about reality, knowledge, backing for knowledge claims, and the value of research I pursue in this thesis? Philosophical pragmatism represents an alternative approach that rejects ontology as the ultimate foundation of knowledge.

Pragmatism

Pragmatism is a philosophical approach initiated by Peirce, James, and Dewey, and later developed by Rorty (1980), Putnam (1990), and Sellars (1997), among others. The term pragmatism prompts a philosophical reflection on the roles knowledge, inquiry, and research play in humans' encounters with our surroundings, rather than a practical approach that favors what is opportune. The pragmatist approach is commonly associated with mixed methods methodology (Morgan, 2007).

Dewey (1920), one of the classical pragmatists, inspired by Darwin's revolution in biology, suggested seeing humans' capacity for knowledge and thinking as an adaption to our environments, rather than a divine capacity to contemplate the universe's harmony. Pragmatists dismiss the question about what reality ultimately is as an essential foundation of knowledge and research. Rather than regarding research as a search for true descriptions of a mind-independent reality, pragmatists focus on inquiry as a search for knowledge about how to cope with obstacles in our surroundings. Knowledge is thus not seen as true and objective representations of an external reality by an agent that disinterestedly observes the world but as instruments and strategies to handle occurring challenges.

Instead of searching for an ultimate foundation for our knowledge, pragmatists focus on how we go along when we search for knowledge that is helpful to us. Typically, inquiries are searches for solutions to the challenges we face. There is a continuity from everyday problem solving towards research. Inquiries of theoretical or philosophical knowledge also represent quests for knowledge located within certain specific contexts. Scientific research is not that different from everyday quests for knowledge; it just entails far more advanced and systematic procedures.

Typically, inquiry and research take place in companionship with other inquirers – in communities of inquiry. Even when someone conduct research alone, the results of the inquiry are approved or contested within a community of fellow inquirers. Rather than seeing inquiry as a search for sentences that are true in the sense that they correctly depict reality, the goal of inquiry is to take us from doubt to stable belief. Since the conundrum about how to account for what truth ultimately is unsolved, pragmatists hold that the result of inquiry is warranted assertability. A correct representation of truth is impossible to account for, yet it is redundant. The goal of our inquiry processes are perceptions that help us cope with our projects, and assertions we can hold as warranted. One of the functions of communities of inquiry is to challenge and scrutinize our perceptions to erase or improve incorrect perceptions, while other perceptions are corroborated.

I present pragmatist philosophy of science as a backdrop for this thesis, as an alternative to the well-established positions from the traditional dichotomies of (post-) positivism and constructivism and quantitative and qualitative methodologies (Morgan, 2014). Pragmatist philosophy of science dismisses the problematic metaphysical discussions about the nature of reality. Rather than ontological mutually incompatible approaches, pragmatism acknowledges arguments from several paradigms (Pearce, 2015). Knowledge based on quantitative methodology may be valued as useful, not because it depicts reality in a more truthful or precise way or has unique access to reality but because it can answer some specific categories of questions and provide knowledge that is more generalizable. Additionally, knowledge based on qualitative methodology may prove useful, for instance, by interpreting meaning and exploring subjects' perceptions. Successful inquiry results in warranted assertions useful when coping with a certain problem. "Problem" should here be understood in the broadest possible sense of the term, and "warranted" is to have endured further scrutiny in a community of inquirers. Further, the pragmatic approach to methodology and philosophy of science seems to coincide better with how scientists actually conduct research. Scrutiny by fellow researchers is the ideal for academic debate and is institutionalized in publication processes as peer review. Reviewers rarely anticipate that researchers apply methodology because they have a certain access to the reality; rather, they consider how well warranted the assertions are.

If I positioned my research in this thesis within a philosophy of science approach based strictly on one or another ontological and epistemological paradigm, claiming a certain perspective to be impermeable, then paradigm-transcending approaches like those I apply in

my thesis would be problematic. This thesis adheres to a pragmatist mixed-method approach by scrutinizing models and operationalizations from multiple theoretical frameworks (Morgan, 2007).The work in my thesis represents a form of theory triangulation by applying arguments from several perspectives.

Distinctions between philosophy of science positions have been reflected in the research field I discuss. For instance, Annand (2019) criticizes the Community of Inquiry framework for being based on a social constructivist paradigm, yet Community of Inquiry researchers employ methods from an objectivist-rational paradigm (similar to what I have termed post-positivism). According to Annand, the Community of Inquiry framework leads to logical (or rather metaphysical) contradictions. According to the pragmatist approach I have presented here (that resonates with the Community of Inquiry framework due to a common affinity for Dewey's thinking), Annand's objection has less relevance.

3.6. Ethics

Guidebooks in educational research and social science commonly focus on ethical obligations towards people researched: informants or participants (see for instance Silverman, 2011). However, other topics may well demand ethical considerations. Several groups may be affected by a research project, not only informants or persons directly involved in the research or mentioned in the reporting of research. Research may afflict the research community, and society in general. In the American Psychological Association's (2010) publication manual, ethical considerations relevant to research publication are discussed under the following headlines: ensuring accuracy of scientific knowledge, protecting rights and welfare of research participants, and protecting intellectual property rights. I have structured my discussion based on affected groups (Tangen, 2014) and according to APA headlines. In the following, I will discuss ethical considerations according to various affected groups, starting with obligations towards the research community (Merton, 1942), then focusing on respect for informants, and finally commenting on possible ethical issues affecting society in general.

Ethical considerations concerning the research community

As researchers, we have obligations towards the ethos of research as a systematic and collaborative quest for better understanding, and towards our fellows in the research community. In this chapter, I already have discussed approaches to ensure the quality of my research: building on previous research, coherence in design, and issues concerning validity.

This has ethical implications in the way that failing to address these ideals may imply failing to fulfill research's potential to contribute to better understanding. Scholarly virtues describe ideal conduct towards colleagues and research community as much they aspire to the ideals of research as an enterprise for better understanding by means of systematic collaborative searches for knowledge. The latter aim is present in ideas open for criticism, scrutiny, and re-testing.

Merton (1942) suggested the acronym CUDOS as a code for the virtues of research. The acronym represents Communism – that knowledge should be free and available for all; Universalism – that "truth-claims […] are subjected to preestablished impersonal criteria" (Merton, 1942, p. 118); Disinterestedness – that the motive for researchers is to test, not to prove, their knowledge claims; and Organized Skepticism – that researchers scrutinize their own, fellow researchers', and societal knowledge. In the next paragraphs, I will comment briefly on two of these virtues: communism and organized skepticism.

Communism and considerations concerning where to publish

Communism (in Merton's use of the term) concerns publication practices, which may have ethical implications. Traditionally, researchers have published their articles in academic journals or in books edited by academic publishers. These have been available in print for readers who buy books or subscribe to journals (commonly handled by university libraries). The publication cost includes printing and physical distribution. Online publication has lowered those expenses. A major concern is that subscriptions are expensive, limiting access for researchers outside well-funded institutions (e.g., those in developing countries). At the same time, academic publishers receive immense revenues. Open Access started as an idealistic movement aiming to make research publications freely available by make most of the affordances of online publishing. Recently, adherence to Open Access has been official policy in Norway and the EU. I submitted Study 1 and 2 to no-cost Open Access journals (known as Gold Open Access). Study 3 has been submitted to the journal *Computers & Education*. Due to an agreement between Norwegian universities and the publisher Elsevier, the universities' subscription fees also serve as a prepayment of fees for Open Access publication (known as Hybrid Open Access).

Academic misconduct or academic dishonesty is to seek unfair academic advantage for oneself or others. This includes plagiarism – using others' work and failing to acknowledge them as well as fraud – using fictitious or manipulated data. Academic misconduct may also include not acknowledging the work of co-authors or honorary co-

authorship (acknowledging someone as co-author for insignificant contributions). A special variant of misconduct towards the research community is to boost publication history by co-operating with predatory publishers and not using peer-review or other established standards for academic publishing. After the publication of Study 1, I regularly receive e-mail invitations to re-publish the text by predatory publishers like Lambert Academic Publishing aka VDM Verlag Dr Müller aka OmniScriptum. Of course, I ignored such invitations. The studies in this thesis are submitted to acknowledged academic journals.

Organized skepticism

Organized skepticism concerns how researchers relate to previous research. The rationale for organized skepticism is not parricide; the rationale is rather to test theories in order to identify those that are corroborated and those that need to be improved. My aim in the thesis is to critically discuss how previous researchers operationalized key concepts in their work. The conclusions I tend towards are that both the Community of Inquiry framework's construct of Cognitive Presence and frameworks based on Toulmin's model fail as operationalizations of the quality of critical thinking and rational argumentation. Employing these models without exposing them for organized skepticism would be to forsake this virtue of research.

Ethical considerations concerning informants

In Study 2, I analyzed students' contributions in OEDs, collected from an introductory course in philosophy at UiT. The discussions were mandatory learning activities and took place in a closed learning management system accessed by students and teachers. In addition to doing research, I worked as an educational counselor and faculty developer. In this role, OEDs were accessible when teachers asked for advice on how to scaffold them. Such data is easily accessible for storage and analyses for educational or administrative purposes. On the other hand, access for such data for research purposes is highly regulated.

As a researcher, I received access to students' contributions in OEDs after the completion of the course. In addition, teachers in the course were interviewed to provide a richer picture of the context. Both students and teachers received information and were asked for consent to use their material for research. NSD (Norwegian Data Protection Services) were notified (see appendix) of the research project since it included collection and storage of both directly identifiable personal data (names) and indirectly identifiable personal data. Immediately after data collection, I deleted names and replaced them with fictitious aliases.

Contributions from students who did not give consent to participate were removed. The material contains no information on sensitive personal data. NSD lists "religious or philosophical beliefs" among special categories of personal data. In this case, the students were asked to reflect on the rather existential topic "the meaning of life." This may include personal experiences and thoughts about sensitive topics. Nevertheless, the students' writings were sober and not characterized by an exposure of personal experiences or existential beliefs. In the publication of the study, I anonymized the students and largely condensed and reformulated their writings. No information that might disclose students' identity of offend their privacy are thus made public. Based on this account, I can claim that this research entails minor ethical risks, and I handled these risks adequately.

Nevertheless, there is always a risk for anyone evaluating one's own conduct self-righteously for ethical appropriateness. Did I fail to recognize my own blind spots thus compromising students' rights in my analysis of their writings? Informants' rights do not concern only privacy and anonymity. Another research ethical virtue is to threat informants and other involved parties with dignity. One of the anonymous reviewers on a previous draft of study 2 characterized my approach as an example of "negative scholarly rationalism" (referring to Rommetveit's phrase), meaning that scholars (in this case, me – the researcher) know the rules, and that nobody else, including the students, should dare to think they do. The reviewer found my evaluation of the students' arguments as "weak", "incomplete", etc. as *besserwisser*-like, not acknowledging that the students were in a learning process and it was thus unreasonable to expect them to master a more profound conception of rational argumentation. The reviewer might be right. I might have set unrealistic standards for the students, disrespecting their efforts and thus their dignity. On the other hand, my intention was never to assess the students' contributions, rather to 1) analyze how a set of categories functioned as analytic tools, and 2) explore what took place in OEDs as a learning activity. My aims were thus focused on the functioning of the analytic tool and the learning activity, not the students' skills. Further, in the discipline of philosophy, evaluating the quality of arguments is not obnoxious; it is the name of the game. My evaluation did to a large degree coincide with some of the students' comments about the discussion. Anyhow, the reviewer's comments were unexpected and reminded me of the possibility for blindness of tendentiousness in my own writing. The comments gave me an opportunity to clarify my arguments and place emphasis on the qualities of the students' writings.

Researching the work of teachers in higher education may also cause certain considerations. Many teachers in higher education are also researchers. In this case, the teachers conducting the course were former colleagues for me and my supervisor (who conducted the interview of the teachers with me). There has been a strong tendency in higher education to consider teaching a private zone. Many teachers report that a colleague has rarely observed them and that discussions about teaching occur in private settings (Roxå & Mårtensson, 2009). Letting a former colleague, now in the role of the researcher, into one's teaching might be scary. I have aimed to avoid that my discussion of the material can be read as an evaluation of their work. Some faculty may even find the use of the term *teachers* instead of *lecturers* or *professors* (*foreleser* is the common Norwegian term) insulting and disrespectful. For me, this is not an attempt to disrespect academic status, but rather signals a student-centered approach to university teaching (Biggs & Tang, 2007; Ramsden, 1992).

University teachers and students can hardly be characterized as vulnerable groups. The minor stress informants possibly experience from being research subjects is likely worth the possible benefits of this study being conducted. Lifting the curtains to see how teaching and learning play out in an OED may faintly offend the privacy of the informants and I consider this risk tolerable.

Ethical considerations concerning society in general

Research may have consequences for society at large. Research has developed knowledge that has led to both catastrophes and solutions for demanding challenges. Considering possible societal consequences of research is obviously an ethical obligation. In a research project like this, the risks for harm are limited. If my arguments are received and accepted, some researchers may become disappointed or annoyed. On the other hand, the potential benefits of my research may also be limited. However, a better understanding of critical thinking and rational argumentation leads to better research and better educational practices.

My PhD scholarship was funded by the University of Tromsø, and thus indirectly by the Norwegian taxpayers. I have spent more the time than allocated for this work, and kind colleagues have enabled me to complete it. My nightmare while completing the work is that anyone would ask me: Is this thesis worth 4 million Norwegian kroner and the extra time you have spent on it? Luckily, I do not have to defend the resources spent on this work. I will return to possible implications and wider applications in the final discussion, chapter 5.

4. Summary of the thesis

In this chapter, I will first recall the research problems that have motivated this thesis and the research questions I developed to investigate them. I will then give a brief overview of the three studies and how they relate to the research questions and sum up my answers to the research questions.

In the introduction, I stated one background research problem in the research field: *How does participation in OEDs enhance students' capacity for critical thinking and rational argumentation?* The problem is urgent for those who are interested in the digitalization of education and the development of critical thinking and rational argumentation. Making use of digital tools in education for interaction among students and development of generic competencies such as critical thinking and rational argumentation, is central to contemporary education and faculty development. I touch briefly upon this background problem in Study 2.

However, the background research problem led me to a methodological problem, which is the main issue in my thesis: *How can students' critical thinking and rational argumentation be analyzed in OEDs?* I pursue this problem by evaluating two of the most influential frameworks in the field, raising two research questions:

1. How is critical thinking analyzed in the Community of Inquiry framework and how adequate is this operationalization?
2. How do frameworks building on Toulmin's argument model analyze quality of argumentation, and how adequate is this way of operationalizing rational argumentation?

In the introduction, I stated that aim of the thesis was to critically evaluate these frameworks, and thereby contribute to a discussion about the quality of research tools and methods in this specific field on OEDs. My evaluation is summarized in my answers to the research questions in this chapter. In chapter 5, I will discuss how my attempt to contribute to a discussion about research tools and methods has been received.

Summary of the three studies

The relation between the studies, the research problems, and research questions, is summarized in Figure 4 in chapter 1. Now, in this summary chapter, I elaborate and present the conclusions I draw. First, I present the three studies:

Study 1

Breivik (2016): Critical Thinking in Online Educational Discussions Measured as Progress through Inquiry Phases: A Discussion of the Cognitive Presence Construct in the Community of Inquiry Framework. *International Journal of E-Learning & Distance Education/Revue internationale du e-learning et la formation à distance, 31(1).*

In Study 1, I approached research question 1 for the thesis by critically evaluating the Cognitive Presence construct from the Community of Inquiry model. The Community of Inquiry framework describes e-learning as three overlapping constructs: teaching presence, social presence, and cognitive presence. Cognitive presence describes critical thinking and is operationalized as progress through phases of an inquiry process. The model is allegedly based on Dewey's account of inquiry as a collaborative enterprise.

Three questions guided my evaluation: First, I traced the philosophical origins of the construct and examined what ideas the framework pursues. Second, guided by the concept of construct validity, I compared the coding scheme with a "minimum conception" of critical thinking. Third, I briefly questioned which empirical findings the framework enables and whether the findings shed light on the model's capability to assess critical thinking.

In the study, I argued that a focus on progress through phases of inquiry does not necessarily emphasize the social aspect of critical thinking, nor does progress through phases in line with Dewey's approach to quality of thinking; rather, I found it lacked construct validity as an operationalization of critical thinking. Finally, low levels of cognitive presence observed by empirical studies were explained by ad hoc hypotheses rather than discussion of the construct's adequacy.

Study 2

Breivik (2020a) Argumentative patterns in students' online discussions in an introductory philosophy course – Micro- and macrostructures of argumentation as analytic tools *Nordic Journal of Digital Literacy, 01-02/2020*

In Study 2, I analyzed transcripts from OEDs in an introductory course in philosophy. The aim of the study was twofold: First, it explored what occurred in online discussions when competence in argumentation were a specific learning objective, analyzed using the categories of the microstructure and macrostructure of argumentation. Second, it discussed how suitable the categories from Toulmin's model are for such analysis.

The students' contribution in the discussions were rich and demonstrated comprehension of the course content, yet my analysis revealed some characteristics of students' argumentation – arguments tend to be unaddressed and incomplete. This analysis

coincides with some of the students' own reflections, and with the actual teachers' overall experience with OED. This way, the study relates to the background research problem about how participation in OEDs contribute to students' development of competency in rational argumentation by showing how such discussions may play out. Even though I can hardly claim that the findings are typical or generalizable, yet they demonstrate both potential and pitfalls when using OEDs as a learning activity.

Further, the Toulmin categories proved to be useful as means to analyze discussion transcripts. Nevertheless, the analysis also showed that coding naturally occurring language (like the students' postings) according to Toulmin's argument categories were cumbersome. In naturally occurring language, discussants rarely structure their arguments according to Toulmin's model. This called for a discussion of how adequate such categories are for analysis of students' arguments. I concluded that analysis by Toulmin-categories provided a profound picture of what took place in the discussions, yet that such analysis missed important qualities of the discussions. Thus, the study demonstrated that presence of Toulmin's argument components used as analytic strategy deserved to be scrutinized. This led to research question 2 for the thesis, about how frameworks building on Toulmin's argument model are used to analyze quality of argumentation, and the adequacy of this this way of operationalizing rational argumentation.

Study 3

Breivik (2020b) Toulmin's Argument Model Used to Analyze Critical Thinking in Online Educational Discussions – An overview and critical discussion. Revised version under review by *Computers and Education*

Several reviews and handbook chapters (Noroozi et al., 2012; Weltzer-Ward, 2011; Wise & Paulus, 2016) claim that frameworks building on Toulmin's argument model are among the most frequently employed approaches to analyze OEDs. In this study, I give an overview of how the model is used in the research field and discuss critically how adequate the model is as a general measure for rational argumentation. Two research questions, which are closely related to research question 2 in the thesis, guide the study: a) How is Toulmin's argument model used in research on online educational discussions? b) How adequate is this model for the analysis of argumentation and learning in online educational discussions? The research questions for this study mirror the overall research problem – *how to analyze students' critical thinking and rational argumentation in online educational discussion?* – and I answer the study's research question a) by presenting and overview of research studies that

have employed Toulmin's categories for analysis of OED. My overview of how Toulmin's model were used showed variation. Most of the research articles employed a simplified version of the model, typically merging the categories of data; warrant; and baking, and the categories of qualifier and rebuttal (see Figure 2 in this extended abstract and Figure 1 and 2 in the study), yet, some studies employed all the categories from the original model. Further, some studies employed only the dimension of arguments' micro-structure (based on Toulmin's model), while other studies employed additional categories (argument's macro-structure, conceptual quality, etc.). Table 3 in the study gives an overview of variations of how Toulmin's model is used. Despite several reviewers have observed frequent use of Toulmin's model in the research field, no previous studies provide an analysis of variation in how the model is used. This study contributes to the research field by providing such an overview of variance on how the model is used.

Further, by analyzing examples of arguments concerning the shape of the earth, thus addressing the study's research question 2, I discuss the adequacy of Toulmin-based frameworks by demonstrating that arguments might contain all Toulmin's argument components, and yet be of weak argumentative quality, and vice versa.

Summed up

Briefly summed up, my answer to research question 1. – How is critical thinking analyzed in the Community of Inquiry framework and how adequate is this operationalization? – is that critical thinking is analyzed as progress through phases of an inquiry process – triggering event; exploration phase; integration, phase; and resolution phase. My evaluation of this operationalization concluded that progress through phases is inadequate as a general approach to analyze students' critical thinking. A discussion may reach final phases of an inquiry process, yet still represent a low level of critical thinking, or contrary, not progress through initial phases and still contain advanced critical thinking. Nuances and backing for my claim are presented and discussed in Study 1.

The short answer to research question 2 – How do frameworks building on Toulmin's argument model analyze quality of argumentation, and how adequate is this way of operationalizing rational argumentation? –is that frameworks using Toulmin's model recognize presence of arguments components as indicator of argument quality: more components present represent higher argument quality. My discussion in study 3 demonstrates that this his approach to analyzing argument quality lacks adequacy as a general measure for rational argumentation. An argument may contain all argument components and

yet be a weak argument, or contrary, an argument may be coded as lacking argument components, yet still represent high argument quality.

I draw conclusions that differ from consensus in the research literature. In the following, I will recapitulate how I arrived at the questions I have worked on and the conclusions stated in the previous paragraphs. My original intention with this thesis was to contribute in research about how digital tools might transform higher education and stimulate students' development of generic competencies like ability to think critically and present and evaluate rational arguments. This is what I describe as the overall research problem for the thesis. One of the first steps for me were to identify a suitable framework to analyze OEDs.

When I first came across the Community of Inquiry framework, I found it promising. Several of its cornerstones coincide with my views. Important is the assumption that digitalization of education should concern more than dissemination of learning material, namely that interaction among students is vital for learning. Further, I agree that critical thinking and rational argumentation are both important learning objectives and stimulates students' learning and understanding. In addition, I share the reading of Dewey that emphasizes the collaborative aspect of inquiry and thus thinking. The framework relates to an established educational approach, namely Lipman's program for Philosophy for Children, with which I had some experience (Breivik & Løkke, 2007).

Despite these common assumptions, I found the frameworks' concept of critical thinking to be problematic. My initial attempts to analyze discussions by the Community of Inquiry coding scheme confused me[14]. So did my attempts to understand the Practical Inquiry model (see Figure 2 in chapter 1). I did not find progress through phases of an inquiry process to be an obvious indicator of critical thinking. This led me to the discussion of the construct validity of the cognitive presence construct. As I concluded in Study 1, the framework was inadequate for my purpose.

My next step, in Study 2, was to test out coding based on the categories from Toulmin's argument model. I believed these categories to be more robust as analytical tools.

---

[14] Other researchers have faced similar confusions using when attemting to employ the Community of Inquiry'sconstruct of Cognitive Presence and the model of Practical Inquiry:

> In that first study we used Garrison, Anderson, and Archer's (2001) practical inquiry model as a theoretical framework to understand the online talk by coding, counting, and tallying up the segments that fit into each phase. I clearly remember how difficult it was to operationalize the theoretical framework into coding categories and the struggle to come to agreement with other analysts on which segment fit into which category. It seemed like it should be completely straightforward, but it was not, and this sparked my interest in methodological innovation around ways to make sense of online talk (Paulus & Wise, 2019, pp. 2-3).

One possible benefit of the Toulmin-based frameworks is that it quite strictly employs the categories from Toulmin's model, without need of any mediating constructs (like the cognitive presence-construct). In addition, Toulmin's model highlights evaluation of how arguments are backed as central. This coincides with a key aspect of critical thinking (see chapter 2). Nevertheless, the analysis revealed that also this framework deserved to be scrutinized thoroughly. In Study 3, I categorized variations in how Toulmin-based frameworks are employed for analysis of OEDs. Further, I questioned whether the framework enables valid inferences about the general quality of students' argumentation and concluded that the framework was inadequate for analyzing general qualities of rational argumentation. This retrospect on the research process and the relation between the studies illustrates that my conclusions contradict my initial assumptions, and thus that my approach has been to test, rather than to prove my initial assumptions.

I have conducted all three studies as single author. Co-authorship might have brought different arguments into my discussion and would surely have improved reliability of literature screening. Nonetheless, I have not conducted the work in total isolation. I have presented lines of reasoning and preliminary drafts from the studies and received feedback from colleagues at conferences and seminars.

5. Closure

In the previous chapter, I summed up the work in this thesis as attempts to answer these research questions:

1. How is critical thinking analyzed in the Community of Inquiry framework and how adequate is this operationalization?
2. How do frameworks building on Toulmin's argument model analyze quality of argumentation, and how adequate is this way of operationalizing rational argumentation?

In this final chapter, I will first comment upon the aim of the thesis – to raise a debate in the research field. My conclusions differ from other researchers' conclusions; therefore, it may be interesting to see whether my studies have sparked any debate. To answer this, I will comment the responses I have received on Study 1. The study concerns research question 1, the Community of Inquiry framework and its Cognitive Presence construct. At the time of writing, Study 2 was recently published while Study 3 is still in the publication process, thus they have not received any responses yet. Therefore, I will only comment briefly here on my evaluation of frameworks applying Toulmin's model and its relation to the field at large.

Next, I will comment on topics related to my research questions that I have not discussed. I do this to relate my research to relevant discussions in the research field and to discuss possible limitations in my thesis.

Finally, I will sketch some wider applications for the claims I hold as the topics I address may be relevant to general discussions about critical thinking and rational argumentation; and teaching and learning in general.

5.1. Have I stirred any debate? Reception of Study 1

Garrison (one of the founders of the Community of inquiry-framework) has in the introduction to his book *Thinking Collaboratively* expressed that in research, there is commonly a risk

> … that humans are inherently selective in seeing and reinforcing existing beliefs. If thinking is to be innovative, there is a need to break out of this cognitive straightjacket and to consider new ideas; to overcome the human bias to confirm and not question currently held perspectives. Experience and evidence is unwittingly selected and interpreted to fit within the individual's personal paradigm. This reluctance to explore conflicting arguments or ideas is well known in science (Kuhn 1962). […] Constructing

> personal meaning without critical feedback can be inherently satisfying but it can also be delusional (Garrison, 2016, p. 2)

This quote beautifully illustrates the general value of raising critical and alternative approaches. In my initial work with the Community of Inquiry framework, I noticed that there was limited critique of the framework, and that critical comments (Jézégou, 2010; Rourke & Anderson, 2004; Rourke & Kanuka, 2007, 2009) were neglected or immunized (Akyol et al., 2009; Garrison, 2013).

My critique concerns only one part of the framework (the Cognitive Presence construct), and I believed that improving parts of the model, or at least considering possible problems, would be interesting for the proponents of the framework. Only a few other commentaries dig into the challenge I address. In 2004, Rourke and Anderson (2004) noted that one explanation for the observed low levels of cognitive presence might be a result of shortcomings in the coding protocol. Later, Ho and Swan (2007) remarked that established approaches to measuring cognitive presence were "less successful" or "yielded disappointing results," thus they suggested an alternative coding scheme (p. 4). Such critical approaches towards the Cognitive Presence construct have not gained much attention.

In the introduction, I stated that my aim with the thesis is to critically evaluate two frameworks and thereby contribute to a discussion about research tools and methods in this specific field on OEDs. According to the Community of Inquiry framework, I raised the argument that one should carefully consider whether the cognitive presence indicators (progress through phases of an inquiry process) really describe what they are intended to describe. Study 1 has received some responses since its publication. I will comment on the reception of study since this gives me an opportunity to reflect on whether I have succeeded in my aim to bring critical evaluation into a scholarly debate.

I submitted the paper to *International Journal of E-Learning and Distance Education*. One of the anonymous reviewers commented: "I find the article's argument persuasive", and further, "I think it may be likely to stir a debate". The second reviewer commented that "[t]his is an important conceptual paper. [It] is extremely relevant and addresses a major issue on online asynchronous discussions." Further, "[the paper gives] a clear explication of the Garrison et al model as well as the illusive concept of critical thinking". I launched the paper into to what I supposed to be a space of debate and critique, based on the Garrison's (2016) quote with which I opened this section. Based on the feedback from the reviewers, I expected some kind of response, and what I first got was silence.

Together with Terry Andersson, Walter Archer, Norm Vaughan, Marti Cleveland-Innes and J.B. Arbaugh, Dr. Randy Garrison is one of the founders of the Community of Inquiry framework and still a leading researcher and proponent. More than a year after my paper was online, Garrison posted a comment to it on "The Community of Inquiry multi-author blog."[15]

Garrison's response to my paper followed two lines of argument: 1) I have misread the Cognitive Presence construct as an operationalization of critical thinking. According to this objection, Garrison claimed that, the Cognitive Presence construct was not intended as an operationalization of critical thinking. Rather, 2) the Cognitive Presence construct addresses other valuable educational experiences, such as "personal reflection and critical discourse in a community of inquiry" (Garrison, 2017).

According to the first of Garrison's objections, that the Cognitive Presence construct is not intended to measure critical thinking, it is illuminating to re-read the seminal paper where Garrison et al. (2001) introduced the construct. Consider the title "Critical Thinking, Cognitive Presence, and Computer Conferencing in Distance Education" which clearly establishes a relation between critical thinking and cognitive presence. Similarly, several quotes in the text reveal a close affinity between the construct of Cognitive Presence and critical thinking; for example, the "quality of cognitive presence, as defined and assessed by the phases of a generalized model of critical thinking" (p. 3), and "[w]e justify this [coding] procedure by noting that higher levels of critical thinking, such as integration and resolution …", (p. 17). Subsequent research conducted by both Garrison himself and a vast number of other researchers demonstrates a close relation between the construct of Cognitive Presence and critical thinking. In my opinion, Garrison's reply that the Cognitive Presence construct is not about critical thinking, seems flawed. Both seminal texts and the reception history illustrate this.

When responding to Garrison's blog post, I addressed his second point – that the Cognitive Presence construct describes other valuable educational experiences, such as "personal reflection and critical discourse in a community of inquiry". The same problem occurs for progress through phases of an inquiry process as indicators for other valuable educational experiences. Both personal reflection and critical discourse may well take place even if the inquiry process does not move beyond initial phases. Conversely, an inquiry

[15] http://www.thecommunityofinquiry.org/editorial5

process may reach the solution phase without much personal reflection and critical discourse taking place. My concern to Garrison and other researchers building on the Community of Inquiry model and the Cognitive Presence construct is that this construct, operationalized as progress through the phases of an inquiry process, lacks adequacy as a measure of deep and meaningful learning, critical thinking, or any other important educational rationale enhanced by collaborative inquiry. My impression after the exchange of views with Garrison is that my argument still holds, and Garrison failed to consider my objections.

Since a few years have passed since the study was published, I have the opportunity to consider other reactions to my text. According to a Google Scholar search performed in April 2020, six academic journal articles (excluding one in Malaysian) and one book chapter cite the study (see Appendix 2). In addition, 11 doctoral theses and one master thesis (all from US and Canada) cite my study. I scanned the journal articles and the book chapter to see if my arguments have stirred any debate.

Disappointingly, four of the journal articles (DiPasquale & Hunter, 2017; Foo & Quek, 2019; Moore, Oliver, & Wang, 2019; Siburian, Corebima, & Saptasari, 2019) that cite Breivik (2016) use my study as backing for the claim that the Community of Inquiry framework is widely accepted or support it in other ways use my paper as support for the CoI-model. This is far from, and even contrary to my main argument in the paper. One article (Heilporn & Lakhal, 2020) and a book chapter written by a leading proponent of the Community of Inquiry model (Swan, 2019) both acknowledge my study as a critique of the model, yet they do not engage in a substantive evaluation of my argument and maintain their positive view of the framework. In contrast, Annand (2019) criticizes the Community of Inquiry framework for contractively mixing social constructivist and objectivist-rational philosophy of science, suggesting that the framework overvalues the collaborative aspect of learning. He cites my paper to claim that progress through phases of an inquiry process may happen even for an isolated individual. This is in line with my argument in the paper. Nevertheless, my argument was not that collaborative learning is overrated. Rather, I claim that the progress through phases of an inquiry process as indicator for successful learning do not capture what it is intended to; namely, learning as an essential collaborative enterprise. Thus, Annand (2019) misreads my critique that the adequacy of the *indicators* is a critique of the adequacy of collaborative approaches to learning.

It is strange to see my text used by other researchers to back claims that are far from, and even contrary to, my argument in the paper. In Study 3 (Breivik, 2020b) I use as example

a line of reasoning that claims to prove that the earth is flat. I hope that no one will cite my study as backing for this idea.

Reading the commentary of Garrison (2017) and other researchers on my study gives me the impression that my critical input has hardly received attention. Thus, I have only succeeded to a limited degree in the aim I stated in the introduction – to contribute in a discussion about research tools. I introduced this section by a quote from Garrison (2016) praising the value of critical approaches, and my observation that critical approaches to the Community of Inquiry framework seem relatively sparse. Stenbom (2018), sympathetic to the Community of Inquiry framework, warns against the risk of that Community of Inquiry research developing into what he characterizes as a "mutual admiration community" (Stenbom, 2018, p. 27). He bases his warning on the observation that a limited number of researchers are involved in a substantive amount of the research publications in the field. Stenbom (2018) reviewed a different branch of the Community of Inquiry research field than my thesis: his review addressed a survey instrument while my research addresses coding schemes for content analysis of discussion transcripts. However, both methodologies build on the same ideas of overlapping presences and progress through phases of inquiry. Thus, his warning is relevant for my argument. This is corroborated by the fact that the researchers he lists (like Garrison; Clevaland-Innes; Ice; Akyol and Shea) have also published extensively in the research branch I discuss.

Study 2 and 3 in relation to the research field

In terms of research question 2 – How do frameworks building on Toulmin's argument model analyze quality of argumentation, and how adequate is this way of operationalizing rational argumentation? – Study 2 and Study 3 have not received any responses yet. However, the research literature on the use of Toulmin's argument categories used for analysis of OED is more diverse. Toulmin's model is commonly held as validated by referring to the model's acknowledged status in philosophy and argumentation theory. Those researchers who built their research on the Toulmin model (see Study 3 for an overview) have certainly concluded that the model is adequate, contrary to my conclusion in this thesis.

Nevertheless, several influential researchers (Andriessen & Baker, 2014; Noroozi et al., 2012; Nussbaum, 2011) have raised objections towards the use of Toulmin's categories as indicators of argumentative quality. Despite this, researchers do not always consider these objections when selecting analytical tools and coding schemes for their analysis. Still,

framework based on the Toulmin model is used (see for instance Lin, 2019 and other researchers cited in my study) and seen as adequate for analyzing argumentative quality.

5.2. Limitations

I described my success above in terms of sparking a debate in the research field as meager. Have I addressed the important issues in the debate? Do I approach the research frontier, or are my analyses and claims dated and of little interest in the contemporary debate? In the subsequent paragraphs, I comment on some the central questions from the research literature I have left out of my discussions.

Relevance and timing

Seminal publications that sparked my interest in this field were published ten to twenty years ago. De Wever et al. (2006) review, which compared several frameworks in the field, and formed the outset for my interest, is not entirely new. Henri (1992) suggested the first framework for analyzing OEDs nearly 30 years ago. At that time, using the Internet for education was a novel idea, and even more innovative was the idea to use the Internet for student-student interaction. In the early decades of the Internet, online forms of communication were rather avant-garde. Pioneers employed e-mail groups (Dysthe, 2001; Marttunen, 1997) and Usenet newsgroups. The possibility of asynchronous yet rapid two-way communication was a radically new affordance, particularly for distant education. The entry of the World Wide Web in the nineties made online communication widespread; Web 2.0, later on, improved usability; and social media and mobile access to the Internet made online conversations ubiquitous. The fascination for the educational affordance of online asynchronous communication may be fading out as new tools and media of communication attract interest. Some of the researchers I address may have left the topic in favor of promising new adaptions of educational technology, like MOOCs, flipped classroom, virtual reality and AI. Perhaps I take up a discussion that has lost significance? However, research studies employing the Community of Inquiry framework or Toulmin's model for analysis of OEDs (Lin, 2019; Liu, Liu, & Lin, 2019) are still published. The frequent use of such frameworks still elicits critical evaluation.

The interest for developing research tools to analyze dialogue in various educational settings is still growing (Erduran, 2018; Erduran et al., 2015; Hennessy, Howe, Mercer, & Vrikki, 2020; Hennessy et al., 2016; Howe, Hennessy, Mercer, Vrikki, & Wheatley, 2019). In

addition, the emphasis on critical thinking and deep learning in curriculum reforms, like the Subject Renewal in Norway, actualizes discussion on how critical thinking may be conceptualized, enhanced, and assessed.

Questions about reliability of frameworks

A topic that has been important in the methodological literature on frameworks for analyzing OEDs is reliability, which concerns whether the same result will be obtained when different researchers code the same material (inter-rater reliability) or when one researcher codes identical material (intra-rater reliability). De Wever et al. (2006) remarked that reliability is a central quality when considering frameworks for analysis of OEDs, yet some of the frameworks they reviewed lacked estimations of reliability. With low or undefined reliability, one can hardly know if analysis is representative of material characteristics or random interpretations. Reliable coding requires precise and observable indicators. When discussing the concept of construct validity (section 3.2), I described the crux of developing coding schemes as identifying indicators that are both representative and observable. My discussion in this thesis focus the representativeness and relevance of indicators.

One challenge I experienced when coding material from OEDs in an introductory philosophy course in Study 2 was that identifying Toulmin's argument components was cumbersome. Due to the implicit style of students' writing, what should be counted as claims or backing, for example, was not evident. In some cases, a possible interpretation is that some of the argument components were given in subsequent posts. Consequently, other coders of the material may have come to different results that those I present in Study 2.

The question I address in this thesis concerns the adequacy; the relevance and representativeness of indicators. To identify observable indicators rely on identifying relevant and representative features of the phenomenon under investigation.

Analytic units and segmentation of data material

Another topic discussed in the field, which I have not addressed in my thesis, concerns what should be coded as the analytic unit: sentences, units within students' posts, single posts, or sequences of several posts? Several researchers have concluded single posts are convenient units for coding.

The choice of analytic unit may have consequences for how critical thinking and rational argumentation (or other constructs) are conceptualized. For instance, using single posts as an analytic unit may highlight individual aspects of learning. If a researcher aims to

highlight collaborative aspects of learning, units stretching over individual contributions may be more appropriate.

Analyzing dialogues based on units either smaller or larger than a single post or utterance may be fruitful. In a study of short in-lecture discussions among students, Ludvigsen, Krumsvik, and Breivik (2020) used whole discussions as analytic units and coded them according to their use of subject terms and productivity (i.e. whether students collaborated to achieve better understanding). The discussants built on and completed each other's arguments. Further both the single discussions and the individual utterances were very short and thus individual utterances as analytical unit would have been difficult to code.

In other analyses, units smaller than a single post or utterance may be more appropriate. In the material analyzed in Study 2, students wrote rather lengthy posts. To analyze this material, analytic units smaller than a single post (e.g., one sentence, one paragraph, or one argument) may be equally appropriate. On the other hand, some arguments in this material stretched beyond one single post; for instance, a rebuttal or qualification was presented in a single post after the main one.

In my thesis, I have not delved into these topics of reliability and analytic units. Discussions of adequacy of indicators are most urgent. Indicators and frameworks may be reliable but inadequate. Even if one chooses small or more extensive units of analysis, the question of indicators' adequacy cannot be dismissed.

5.3. Wider applications

Critical thinking and rational argumentation as educational virtues receive growing interest. Digitalization of education has also been on the agenda over the past decades. In chapter 2, I discussed these topics as the background for my research. The specific topic I have been scrutinizing in this thesis – how to operationalize and analyze critical thinking and rational argumentation in OEDs – is a rather tiny stroke on a greater canvas. Yet, the discussions I take on in the thesis relate to the wider landscape of contemporary education and educational research. The increasing emphasis on critical thinking and rational argumentation is commonly related to other generic competencies, such as problem solving and deep learning. In a Norwegian context, the curriculum reform *Fagfornyelsen* (the Subject Renewal) (Kunnskapsdepartementet, 2016) highlights critical thinking as a core value for education. In chapter 2, I referred to research that claimed many educators lack a clear conception of critical thinking (Paul et al., 1997), and this seems characteristic of the field.

Research on both assessment and formative feedback (see for instance Hattie & Timperley, 2007) and teaching and learning in general (see for instance Anderson & Krathwohl, 2001; Biggs & Tang, 2007; Hattie, 2015) states that clear and explicit learning objectives and criteria for success promote student learning. Without a clear conception, teaching tends to rely on osmosis or serendipity, and the students must figure out the key characteristics of critical thinking themselves.

Osmosis describes a physical process where a solvent move through a semi-permeable membrane, from a higher concentration to a lower concentration, like saltwater penetrating a cell. As a metaphor for learning, osmosis takes place as a result of a learner being situated in a salient milieu or through implicit model learning. Serendipity describes a similar implicit approach to learning, when a learner stumbles upon an insight (or something worthwhile) by chance or luck rather than intent or planning. These approaches resemble what Ennis (1989) has labelled the immersion approach to teaching critical thinking. When teaching critical thinking and rational argumentation without explicating the key features for the learners, learners risk missing key aspects (e.g., evaluation of claims' acceptability or relevance), or wrongly interpreted less central aspects as crucial (for instance being "critical" and oriented towards dispute).

By thematizing concepts like critical thinking and rational argumentation, I hope to inspire educators to clarify for their students how critical thinking and rational argumentation may be understood as learning objectives and assessment criteria. A thorough conceptualization of critical thinking and rational arguments may improve educational practices. A proper understanding of these concepts is essential not only for research purposes but when planning and conducting education and assessing students' learning.

References

AERA, NCME, N. C. o. M. i. E., & APA, t. (2014). *Standards for Educational and Psychological Testing*. Washington, D.C: American Educational Research Association.

Akyol, Z., Arbaugh, J. B., Cleveland-Innes, M., Garrison, D. R., Ice, P., Richardson, J. C., & Swan, K. (2009). A Response to the Review of the Community of Inquiry Framework. *Journal of Distance Education, 23*(2), 123-135. http://www.jofde.ca/index.php/jde/article/view/630/884.

Anderson, L. W., & Krathwohl, D. R. (2001). *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives*. New York: Longman.

Andriessen, J., & Baker, M. (2014). Arguing to Learn. In R. K. Sawyer (Ed.), *The Cambridge Handbook of the Learning Sciences* (2 ed., pp. 439-460). Cambridge: Cambridge University Press.

Anfinsen, R., & Laugerud, S. (2007). *Nettdiskusjoner: Faglig Dialog Eller "God Dag, Mann! - Økseskaft"?* (978-82-91308-45-6). In: Myklebost &Skared (ed.) *Om remidiering av undervisning - og læring gjennom samarbeid. Refleksjoner over prosjekterfaringer.* Norgesuniversitetets skriftserie 2/2007.

Arbaugh, J. B., Cleveland-Innes, M., Diaz, S. R., Garrison, D. R., Ice, P., Richardson, J. C., & Swan, K. P. (2008). Developing a Community of Inquiry Instrument: Testing a Measure of the Community of Inquiry Framework Using a Multi-Institutional Sample. *The Internet and Higher Education, 11*(3–4), 133-136. R doi:http://dx.doi.org/10.1016/j.iheduc.2008.06.003

Arum, R., & Roksa, J. (2011). *Academically Adrift: Limited Learning on College Campuses*: University of Chicago Press.

Bhaskar, R. (2008). *A Realist Theory of Science*: Taylor & Francis.

Biasutti, M. (2017). A Coding Scheme to Analyse the Online Asynchronous Discussion Forums of University Students. *Technology, Pedagogy and Education, 26*(5), 601-615. doi:10.1080/1475939X.2017.1365753

Biggs, J. B., & Tang, C. S.-K. (2007). *Teaching for Quality Learning at University : What the Student Does*. Maidenhead: McGraw-Hill/Society for Research into Higher Education & Open University Press.

Blair, J. A. (2012). Relevance, Acceptability and Sufficiency Today. In C. W. Tindale (Ed.), *Groundwork in the Theory of Argumentation: Selected Papers of J. Anthony Blair* (pp. 87-100). Dordrecht: Springer Netherlands.

Boote, D. N., & Beile, P. (2005). Scholars before Researchers: On the Centrality of the Dissertation Literature Review in Research Preparation. *Educational researcher, 34*(6), 3-15.

Bostrom, R. N. (2003). Theories, Data, and Communication Research. *Communication Monographs, 70*(4), 275-294. doi:10.1080/0363775032000179106

Boyer, E. L. (1990). *Scholarship Reconsidered: Priorities of the Professoriate*: Princeton University Press, 3175 Princeton Pike, Lawrenceville.

Breivik, J. (2016). Critical Thinking in Online Educational Discussions Measured as Progress through Inquiry Phases: A Discussion of the Cognitive Presence Construct in the Community of Inquiry Framework. *International Journal of E-Learning & Distance Education/Revue internationale du e-learning et la formation à distance, 31*(1).

Breivik, J. (2020a). Argumentative Patterns in Students' Online Discussionsin an Introductory Philosophy Course, Micro- and Macrostructures of Argumentation as Analytic Tools. *Nordic Journal of Digital Literacy, Vol. 15*(No. 1-2020), 8-23.

Breivik, J. (2020b). Toulmin's Argument Model Used to Analyze Critical Thinking in Online Educational Discussions - an Overview and Critical Evaluation. *Under review by Computers & Education*.

Breivik, J., & Løkke, H. (2007). *Filosofi I Skolen - En Kunnskapsoversikt*. Kunnskapsdepartement, Oslo. Retrieved from https://www.regjeringen.no/globalassets/upload/kd/vedlegg/rapporter/filosofi-i-skole.pdf

Bøyum, S. (2006). The Legitimacy of Critical Thinking: Political Liberalism and Compulsory Education. *Thinking: The Journal of Philosophy for Children, 18*(1), 31-39.

Cahill, A., & Bloch-Schulman, S. (2012). Argumentation Step-by-Step: Learning Critical Thinking through Deliberate Practice. *Teach Philos., 35*(1), 41-62.

Cho, K. L., & Jonassen, D. H. (2002). The Effects of Argumentation Scaffolds on Argumentation and Problem Solving. *Etr&D-Educational Technology Research and Development, 50*(3), 5-22. doi:Doi 10.1007/Bf02505022

Clarà, M., & Mauri, T. (2010). Toward a Dialectic Relation between the Results in Cscl: Three Critical Methodological Aspects of Content Analysis Schemes. *International Journal of Computer-Supported Collaborative Learning, 5*(1), 117-136. Retrieved from http://dx.doi.org/10.1007/s11412-009-9078-4. doi:10.1007/s11412-009-9078-4

Clark, D., Sampson, V., Weinberger, A., & Erkens, G. (2007). Analytic Frameworks for Assessing Dialogic Argumentation in Online Learning Environments. *Educational Psychology Review, 19*(3), 343-374. Retrieved from http://dx.doi.org/10.1007/s10648-007-9050-7. doi:10.1007/s10648-007-9050-7

Clark, D. B., & Sampson, V. D. (2008). Assessing Dialogic Argumentation in Online Environments to Relate Structure, Grounds, and Conceptual Quality. *Journal of Research in Science Teaching, 45*(3), 293-321. doi:10.1002/tea.20216

Creswell, J. W. (2014a). *Educational Research : Planning, Conducting, and Evaluating Quantitative and Qualitative Research* (4th ed., New international ed. ed.). Harlow: Pearson.

Creswell, J. W. (2014b). *Research Design : Qualitative, Quantitative, and Mixed Methods Approaches* (4th ed.; International student ed. ed.). Los Angeles, Calif: SAGE.

Crotty, M. (1998). *The Foundations of Social Research : Meaning and Perspective in the Research Process*. London: Sage.
Dannelsesutvalget. (2009). *Kunnskap Og Dannelse Foran Et Nytt Århundre. Innstilling Fra Dannelsesutvalget for Høyere Utdanning*. Retrieved from https://w2.uib.no/filearchive/innstilling-dannelsesutvalget.pdf
Davies, M., & Barnett, R. (2015). Introduction. In M. Davies & R. Barnett (Eds.), *The Palgrave Handbook of Critical Thinking in Higher Education*.
De Wever, B., Schellens, T., Valcke, M., & Van Keer, H. (2006). Content Analysis Schemes to Analyze Transcripts of Online Asynchronous Discussion Groups: A Review. *Computers & Education, 46*(1), 6-28. doi:http://dx.doi.org/10.1016/j.compedu.2005.04.005
Dewey, J. (1920). *Reconstruction in Philosophy*: Henry Holt and Company.
Dewey, J. (1933/1986). How We Think (Rev. Ed.) In *The Later Works: 1925-1953 : Vol. 8 : 1933*. Carbondale, Ill: Southern Illinois University Press.
Dewey, J. (1997). *How We Think*. Mineola, N.Y.: Dover Publ.
Dewey, J. (2007). *Democracy and Education : An Introduction to the Philosophy of Education*. [Sioux Falls]: NuVision.
DiPasquale, J., & Hunter, W. (2017). Critical Thinking in Asynchronous Online Discussions: A Systematic Review. *Canadian Journal of Learning and Technology/La revue canadienne de l'apprentissage et de la technologie, 43*(2).
Dysthe, O. (2001). Dialogperspektiv På Elektroniske Diskusjonar. In O. Dysthe (Ed.), *Dialgo Samspel Og Læring*. Oslo: Abstrakt forlag.
Ennis, R. H. (1989). Critical Thinking and Subject Specificity: Clarification and Needed Research. *Educational researcher, 18*(3), 4-10.
Ennis, R. H. (1991/2015). Critical Thinking: A Streamlined Conception. In M. Davies & R. Barnett (Eds.), *The Palgrave Handbook of Critical Thinking in Higher Education* (pp. 31-47). New York: Palgrave Macmillan US.
Erduran, S. (2018). Toulmin's Argument Pattern as a "Horizon of Possibilities" in the Study of Argumentation in Science Education. *Cultural Studies of Science Education*, 1-9. doi:10.1007/s11422-017-9847-8
Erduran, S., Ozdem, Y., & Park, J.-Y. (2015). Research Trends on Argumentation in Science Education: A Journal Content Analysis from 1998–2014. *International Journal of STEM Education, 2*(1), 5. Retrieved from https://doi.org/10.1186/s40594-015-0020-1. doi:10.1186/s40594-015-0020-1
Facione, P. A. (1990). Critical Thinking: A Statement of Expert Consensus for Purposes of Educational Assessment and Instruction. Research Findings and Recommendations.
Foo, S. Y., & Quek, C. L. (2019). Developing Students' Critical Thinking through Asynchronous Online Discussions: A Literature Review. *Malaysian Online Journal of Educational Technology, 7*(2), 37-58.
Frankena, W. K. (1965). *Three Historical Philosophies of Education: Aristotle, Kant, Dewey. [by] William K. Frankena*. Chicago: Chicago, Scott, Foresman.

Garrison, Anderson, T., & Archer, W. (1999). Critical Inquiry in a Text-Based Environment: Computer Conferencing in Higher Education. *The Internet and Higher Education, 2*(2-3), 87-105. doi:10.1016/s1096-7516(00)00016-6

Garrison, D. R. (2011). *E-Learning in the 21st Century*. New York: Routledge.

Garrison, D. R. (2013). Theoretical Foundations and Epistemological Insights of the Community of Inquiry. In Z. Akyol & D. R. Garrison (Eds.), *Educational Communities of Inquiry Theoretical Framework, Research and Practice*. Hershey: Information scinece Reefernce, IGI Global.

Garrison, D. R. (2016). *Thinking Collaboratively: Learning in a Community of Inquiry*. New York: Routledge.

Garrison, D. R. (2017). Cognitive Presence and Critical Thinking. Retrieved from http://www.thecommunityofinquiry.org/

Garrison, D. R., Anderson, T., & Archer, W. (2001). Critical Thinking, Cognitive Presence, and Computer Conferencing in Distance Education. *American Journal of distance education, 15*(1), 7-23.

Gerbic, P., & Stacey, E. (2005). A Purposive Approach to Content Analysis: Designing Analytical Frameworks. *The Internet and Higher Education, 8*(1), 45-59. doi:http://dx.doi.org/10.1016/j.iheduc.2004.12.003

Gettier, E. L. (1963). Is Justified True Belief Knowledge? *Analysis, 23*(6), 121-123.

Gough, D., Thomas, J., & Oliver, S. (2012). Clarifying Differences between Review Designs and Methods. *Systematic Reviews, 1*(1), 28. doi:10.1186/2046-4053-1-28

Grant, M. J., & Booth, A. (2009). A Typology of Reviews: An Analysis of 14 Review Types and Associated Methodologies. In (Vol. 26, pp. 91-108). Oxford, UK.

Hattie, J. (2015). The Applicability of Visible Learning to Higher Education. *Scholarship of Teaching and Learning in Psychology, 1*(1), 79.

Hattie, J., & Timperley, H. (2007). The Power of Feedback. *Review of Educational Research, 77*(1), 81-112.

Hegelund, S., & Kock, C. (1999). *Macro-Toulmin: The Argument Model as Structural Guideline in Academic Writing*. Paper presented at the The Ontario Society for the Study of Argumentation.

Heilporn, G., & Lakhal, S. (2020). Investigating the Reliability and Validity of the Community of Inquiry Framework: An Analysis of Categories within Each Presence. *Computers & Education, 145*, 103712.

Hennessy, S., Howe, C., Mercer, N., & Vrikki, M. (2020). Coding Classroom Dialogue: Methodological Considerations for Researchers. *Learning, Culture and Social Interaction, 25*, 100404.

Hennessy, S., Rojas-Drummond, S., Higham, R., Márquez, A. M., Maine, F., Ríos, R. M., . . . Barrera, M. J. (2016). Developing a Coding Scheme for Analysing Classroom Dialogue across Educational Contexts. *Learning, Culture and Social Interaction, 9*, 16-44. doi:https://doi.org/10.1016/j.lcsi.2015.12.001

Henri, F. (1992). Computer Conferencing and Content Analysis. In *Collaborative Learning through Computer Conferencing* (pp. 117-136): Springer.

Hillen, S. A. (2014). The Role of Discussion Boards in E-Collaborative Learning Environments (Cscl)–What Kind of Support Can They Provide?-a Conceptual Discussion and a Qualitative Case Study. *Nordic Journal of Digital Literacy, 9*(02), 128-147.

Howe, C., Hennessy, S., Mercer, N., Vrikki, M., & Wheatley, L. (2019). Teacher–Student Dialogue During Classroom Teaching: Does It Really Impact on Student Outcomes? *Journal of the Learning Sciences*, 1-51.

Inch, E. S., & Warnick, B. (2011). *Critical Thinking and Communication: The Use of Reason in Argument*: Allyn & Bacon.

Jézégou, A. (2010). Community of Inquiry in E-Learning: A Critical Analysis of the Garrison and Anderson Model. *Journal of Distance Education, 24*(3), 18-27.

Johnson, R. H., & Blair, J. A. (2006). *Logical Self-Defense*: Idea.

Jonassen, D., & Kim, B. (2010). Arguing to Learn and Learning to Argue: Design Justifications and Guidelines. *Educational Technology Research and Development, 58*(4), 439-457. doi:10.1007/s11423-009-9143-8

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why Minimal Guidance During Instruction Does Not Work: An Analysis of the Failure of Constructivist, Discovery, Problem-Based, Experiential, and Inquiry-Based Teaching. *Educational Psychologist, 41*(2), 75-86. doi:10.1207/s15326985ep4102_1

Kleven, T. A. (2008). Validity and Validatio in Qualitative and Quantitative Research. *Nordic Studies in Education, 28*(03).

Kleven, T. A., & Hjardemaal, F. (2018). *Innføring I Pedagogisk Forskningsmetode : En Hjelp Til Kritisk Tolking Og Vurdering* (3. utg. ed.). Bergen: Fagbokforl.

Krumsvik, R. J. (2016). *En Doktorgradsutdanning I Endring : Et Fokus På Den Artikkelbaserte Ph.D.-Avhandlingen*. Bergen: Fagbokforl.

Kunnskapsdepartementet. (2016). *Fag – Fordypning – Forståelse : En Fornyelse Av Kunnskapsløftet (Meld. St. Nr 28(2015-2016)*. Oslo: Departementet

Kunnskapsdepartementet. (2017). *Digitaliseringsstrategi for Universitets- Og Høyskolesektoren 2017-2021*. Retrieved from https://www.regjeringen.no/contentassets/779c0783ffee461b88451b9ab71d5f51/no/pdfs/digitaliseringsstrategi-for-universitets--og-hoysk.pdf

Kvale, S., & Brinkmann, S. (2009). *Interviews : Learning the Craft of Qualitative Research Interviewing* (2nd ed. ed.). Los Angeles, Calif: Sage.

Kvernbekk, T. (2012). Argumentation in Theory and Practice: Gap or Equilibrium? *Informal logic, 32*(3), 288-305.

Leitao, S. (2000). The Potential of Argument in Knowledge Building. *Human Development, 43*(6), 332-360.

Lin, Y.-R. (2019). Student Positions and Web-Based Argumentation with the Support of the Six Thinking Hats. *Computers & Education, 139*, 191-206.

Linell, P. (2009). *Rethinking Language, Mind, and World Dialogically : Interactional and Contextual Theories of Human Sense-Making*. Charlotte, N.C: Information Age Publ.

Lipman, M. (2003). *Thinking in Education* (2nd ed. ed.). Cambridge: Cambridge University Press.
Lipman, M., Sharp, A. M., & Oscanyan, F. S. (1980). *Philosophy in the Classroom*. Philadelphia: Temple University Press.
Liu, Q.-T., Liu, B.-W., & Lin, Y.-R. (2019). The Influence of Prior Knowledge and Collaborative Online Learning Environment on Students' Argumentation in Descriptive and Theoretical Scientific Concept. *International Journal of Science Education, 41*(2), 165-187. doi:10.1080/09500693.2018.1545100
Loncar, M., Barrett, N. E., & Liu, G.-Z. (2014). Towards the Refinement of Forum and Asynchronous Online Discussion in Educational Contexts Worldwide: Trends and Investigative Approaches within a Dominant Research Paradigm. *Computers & Education, 73*(0), 93-110. doi:http://dx.doi.org/10.1016/j.compedu.2013.12.007
Ludvigsen, K., Krumsvik, R. J., & Breivik, J. (2020). Behind the Scenes: Unpacking Student Discussion and Critical Reflection in Lectures. *British Journal of Educational Technology*.
Låg, T., & Sæle, R. G. (2019). Does the Flipped Classroom Improve Student Learning and Satisfaction? A Systematic Review and Meta-Analysis. *AERA Open, 5*(3), 2332858419870489.
Marton, F., & Säljö, R. (1976). On Qualitative Differences in Learning: I - Outcome and Process. *British Journal of Educational Psychology, 46*(1), 4-11. doi:10.1111/j.2044-8279.1976.tb02980.x
Martono, F., & Salam, U. (2017). Students' Learning in Asynchronous Discussion Forums: A Meta-Analysis. *International Journal of Information and Communication Technology Education (IJICTE), 13*(1), 48-60.
Marttunen, M. (1997). Electronic Mail as a Pedagogical Delivery System: An Analysis of the Learning of Argumentation. *Research in Higher Education, 38*(3), 345-363.
Maxwell, J. A. (2006). Literature Reviews of, and for, Educational Research: A Commentary on Boote and Beile's "Scholars before Researchers". *Educational researcher, 35*(9), 28-31.
Maxwell, J. A. (2013). *Qualitative Research Design : An Interactive Approach* (3rd ed. ed. Vol. 41). Los Angeles: Sage.
Mayer, R. E. (2008). Applying the Science of Learning: Evidence-Based Principles for the Design of Multimedia Instruction. *American Psychologist, 63*(8), 760.
McLoughlin, C., & Lee, M. J. W. (2010). Pedagogy 2.0: Critical Challenges and Responses to Web 2.0 and Social Software in Tertiary Teaching. In *Web 2.0-Based E-Learning: Applying Social Informatics for Tertiary Teaching* (pp. 43-69): IGI Global.
McPeck, J. E. (1990). Critical Thinking and Subject Specificity: A Reply to Ennis. *Educational researcher, 19*(4), 10-12.
Merton, R. K. (1942). A Note on Science and Democracy'. *Journal of Legal and Political Sociology, 1*, 115-126.

Messick, S. (1995). Validity of Psychological Assessment: Validation of Inferences from Persons' Responses and Performances as Scientific Inquiry into Score Meaning. *American Psychologist, 50*(9), 741.
Moore, M. G. (1989). Three Types of Interaction. In: Taylor & Francis.
Moore, R. L., Oliver, K. M., & Wang, C. (2019). Setting the Pace: Examining Cognitive Processing in Mooc Discussion Forums with Automatic Text Analysis. *Interactive Learning Environments*, 1-15.
Moore , T. (2004). The Critical Thinking Debate: How General Are General Thinking Skills? *Higher Education Research & Development, 23*(1), 3-18. doi:10.1080/0729436032000168469
Moore, T. (2011a). Critical Thinking and Disciplinary Thinking: A Continuing Debate. *Higher Education Research & Development, 30*(3), 261-274.
Moore, T. J. (2011b). Critical Thinking: Seven Definitions in Search of a Concept. *Studies in Higher Education, 38*(4), 506-522. doi:10.1080/03075079.2011.586995
Morgan, D. L. (2007). Paradigms Lost and Pragmatism Regained: Methodological Implications of Combining Qualitative and Quantitative Methods. *Journal of mixed methods research, 1*(1), 48-76.
Morgan, D. L. (2014). Pragmatism as a Paradigm for Social Research. *Qualitative inquiry, 20*(8), 1045-1053.
Naidu, S., & Järvelä, S. (2006). Analyzing Cmc Content for What? *Computers & Education, 46*(1), 96-103. doi:http://dx.doi.org/10.1016/j.compedu.2005.04.001
Neuendorf, K. A. (2002). *The Content Analysis Guidebook*. Thousand Oaks, Calif: Sage.
Noroozi, O., Weinberger, A., Biemans, H. J. A., Mulder, M., & Chizari, M. (2012). Argumentation-Based Computer Supported Collaborative Learning (Abcscl): A Synthesis of 15 Years of Research. *Educational Research Review, 7*(2), 79-106. doi:http://dx.doi.org/10.1016/j.edurev.2011.11.006
NOU 2015:8. (2015). *Fremtidens Skole : Fornyelse Av Fag Og Kompetanser : Utredning Fra Et Utvalg Oppnevnt Ved Kongelig Resolusjon 21. Juni 2013 : Avgitt Til Kunnskapsdepartementet 15. Juni 2015*. (9788258312397). Oslo: Departementenes sikkerhets- og serviceorganisasjon, Informasjonsforvaltning
Nussbaum, E. M. (2011). Argumentation, Dialogue Theory, and Probability Modeling: Alternative Frameworks for Argumentation Research in Education. *Educational Psychologist, 46*(2), 84-106.
Næss, A. (1975). *En Del Elementære Logiske Emner: Bokmålsutgave*. Oslo: Universitetsforl.
Næss, A. (2005). *The Selected Works of Arne Naess : Vol. 1 : Interpretation and Preciseness : A Contribution to the Theory of Communication* (Vol. Vol. 1). Dordrecht: Springer.
Osborne, J. (2010). Arguing to Learn in Science: The Role of Collaborative, Critical Discourse. *Science, 328*(5977), 463-466.

Paul, R. W., Elder, L., & Bartell, T. (1997). California Teacher Preparation for Instruction in Critical Thinking: Research Findings and Policy Recommendations.

Paulus, T. M., & Wise, A. F. (2019). *Looking for Insight, Transformation, and Learning in Online Talk*. Milton: Taylor & Francis Group, 2019. Web.

Pearce, L. D. (2015). Thinking Outside the Q Boxes: Further Motivating a Mixed Research Perspective. *Oxford handbook of multimethod and mixed methods research inquiry*, 42-57.

Pedró, F. (2012). *Connected Minds: Technology and Today's Learners*: OECD.

Poce, A., Corcione, L., & Iovine, A. (2012). Content Analysis and Critical Thinking. An Assessment Study. *Cadmo*.

Putnam, H. (1990). *Realism with a Human Face*. Cambridge, Mass: Harvard University Press.

Ramsden, P. (1992). *Learning to Teach in Higher Education*. London: Routledge.

Rorty, R. (1980). *Philosophy and the Mirror of Nature*. Oxford: Blackwell.

Rourke, L., & Anderson, T. (2004). Validity in Quantitative Content Analysis. *Educational Technology Research and Development, 52*(1), 5-18. doi:10.1007/bf02504769

Rourke, L., & Kanuka, H. (2007). Barriers to Online Critical Discourse. *International Journal of Computer-Supported Collaborative Learning, 2*(1), 105-126. doi:10.1007/s11412-007-9007-3

Rourke, L., & Kanuka, H. (2009). Learning in Communities of Inquiry: A Review of the Literature. *Journal of Distance Education, 23*(1), 19-48.

Roxå, T., & Mårtensson, K. (2009). Significant Conversations and Significant Networks–Exploring the Backstage of the Teaching Arena. *Studies in Higher Education, 34*(5), 547-559.

Salmon, G. (2002). *E-Tivities : The Key to Active Online Learning*. London: Kogan Page.

Salmon, G. (2004). *E-Moderating : The Key to Teaching and Learning Online*. London: RoutledgeFalmer.

Salmon, G. (2013). *E-Tivities: The Key to Active Online Learning*. New York: Routledge.

Schindler, L., & Burkholder Jr, G. J. (2014). Instructional Design and Facilitation Approaches That Promote Critical Thinking in Asynchronous Online Discussions: A Review of the Literature. *Higher Learning Research Communications. Vol. 4.*

Schneider, M., & Preckel, F. (2017). Variables Associated with Achievement in Higher Education: A Systematic Review of Meta-Analyses. *Psychological bulletin, 143*(6), 565.

Sellars, W. (1997). *Empiricism and the Philosophy of Mind*. Cambridge, Mass: Harvard University Press.

Selwyn, N. (2016). *Education and Technology : Key Issues and Debates* (Second edition. ed.). London: Bloomsbury.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Boston: Houghton Mifflin.

Shea, P. (2010). A Re-Examination of the Community of Inquiry Framework: Social Network and Content Analysis. *The Internet and Higher Education, 13*(1), 10.

Siburian, J., Corebima, A. D., & Saptasari, M. (2019). The Correlation between Critical and Creative Thinking Skills on Cognitive Learning Results. *Eurasian Journal of Educational Research, 19*(81), 99-114.

Siegel, H. (1988/2013). *Educating Reason*: Routledge.

Siegel, H. (2012). Education as Initiation into the Space of Reasons. *Theory and Research in Education, 10*(2), 191-202. doi:10.1177/1477878512446542

Silverman, D. (2011). *Interpreting Qualitative Data: A Guide to the Principles of Qualitative Research* (4th ed. ed.). Los Angeles: SAGE.

Solberg, M. (2011). Educating the Citizen of Academia Online? *The International Review of Research in Open and Distributed Learning, 12*(4), 77-87.

Stenbom, S. (2018). A Systematic Review of the Community of Inquiry Survey. *The Internet and Higher Education, 39*, 22-32.

Swan, K. (2019). Social Construction of Knowledge and the Community of Inquiry Framework. In *Open and Distance Education Theory Revisited* (pp. 57-65): Springer.

Tangen, R. (2014). Balancing Ethics and Quality in Educational Research—the Ethical Matrix Method. *Scandinavian Journal of Educational Research, 58*(6), 678-694.

Thomas, J. (2013). Exploring the Use of Asynchronous Online Discussion in Health Care Education: A Literature Review. *Computers & Education, 69*(0), 199-215. doi:http://dx.doi.org/10.1016/j.compedu.2013.07.005

Toulmin, S. (2003/1958). *The Uses of Argument*. Cambridge: Cambridge University Press.

Toulmin, S., Rieke, R., & Janik, A. (1984). *An Introduction to Reasoning* (2nd ed.). New York: Macmillan.

Twining, P., Heller, R. S., Nussbaum, M., & Tsai, C.-C. (2017). Some Guidance on Conducting and Reporting Qualitative Studies. In: *Computers & Education, vol. 106. A1-A9.*

Uzuner, S. (2007). Educationally Valuable Talk: A New Concept for Determining the Quality of Online Conversations. *Journal of Online Learning and Teaching, 3*(4), 400-410.

Weinberger, A., & Fischer, F. (2006). A Framework to Analyze Argumentative Knowledge Construction in Computer-Supported Collaborative Learning. *Computers & Education, 46*(1), 71-95. doi:http://dx.doi.org/10.1016/j.compedu.2005.04.003

Wells, R. S., Kolek, E. A., Williams, E. A., & Saunders, D. B. (2015). "How We Know What We Know": A Systematic Comparison of Research Methods Employed in Higher Education Journals, 1996—2000 V. 2006—2010. *The Journal of Higher Education, 86*(2), 171-198.

Weltzer-Ward, L. (2011). Content Analysis Coding Schemes for Online Asynchronous Discussion. *Campus-Wide Information Systems, 28*(1), 56-74.

Wise, A. F., & Paulus, T. M. (2016). Analyzing Learning in Online Discussions. *The SAGE Handbook of E-learning Research*, 270.

Yang, Y.-T. (2008). A Catalyst for Teaching Critical Thinking in a Large University Class in Taiwan: Asynchronous Online Discussions with the Facilitation of Teaching Assistants. *Educational Technology Research and Development, 56*(3), 241-264. doi:10.1007/s11423-007-9054-5

Appendices

1. Letter from NSD
2. Table over citations to Study 1

Norsk samfunnsvitenskapelig datatjeneste AS
NORWEGIAN SOCIAL SCIENCE DATA SERVICES

Harald Hårfagres gate 29
N-5007 Bergen
Norway
Tel: +47-55 58 21 17
Fax: +47-55 58 96 50
nsd@nsd.uib.no
www.nsd.uib.no
Org.nr. 985 321 884

Jens Breivik
Institutt for lærerutdanning og pedagogikk
Universitetet i Tromsø
Mellomveien 110
9037 TROMSØ

Vår dato: 04.12.2012 Vår ref:31997 / 3 / IB Deres dato: Deres ref:

TILBAKEMELDING PÅ MELDING OM BEHANDLING AV PERSONOPPLYSNINGER

Vi viser til melding om behandling av personopplysninger, mottatt 01.11.2012. All nødvendig informasjon om prosjektet forelå i sin helhet 22.11.2012. Meldingen gjelder prosjektet:

31997	*Fleksible førstesemesterstudier*
Behandlingsansvarlig	*Universitetet i Tromsø, ved institusjonens øverste leder*
Daglig ansvarlig	*Jens Breivik*

Personvernombudet har vurdert prosjektet og finner at behandlingen av personopplysninger er meldepliktig i henhold til personopplysningsloven § 31. Behandlingen tilfredsstiller kravene i personopplysningsloven.

Personvernombudets vurdering forutsetter at prosjektet gjennomføres i tråd med opplysningene gitt i meldeskjemaet, korrespondanse med ombudet, eventuelle kommentarer samt personopplysningsloven og helseregisterloven med forskrifter. Behandlingen av personopplysninger kan settes i gang.

Det gjøres oppmerksom på at det skal gis ny melding dersom behandlingen endres i forhold til de opplysninger som ligger til grunn for personvernombudets vurdering. Endringsmeldinger gis via et eget skjema, http://www.nsd.uib.no/personvern/forsk_stud/skjema.html. Det skal også gis melding etter tre år dersom prosjektet fortsatt pågår. Meldinger skal skje skriftlig til ombudet.

Personvernombudet har lagt ut opplysninger om prosjektet i en offentlig database, http://pvo.nsd.no/prosjekt.

Personvernombudet vil ved prosjektets avslutning, 01.12.2015, rette en henvendelse angående status for behandlingen av personopplysninger.

Vennlig hilsen

Vigdis Namtvedt Kvalheim

Inga Brautaset

Inga Brautaset tlf: 55 58 26 35
Vedlegg: Prosjektvurdering

Table 2

Appendix. Citations of Study 1, Breivik (2016)

Genre	**Author(s)**	**Year**	**Title**	**Publisher**
Book chapter	Swan	2019	Social Construction of Knowledge and the Community of Inquiry Framework	Springer
Journal article	DiPasquale, & Hunter	2017	Critical Thinking in Asynchronous Online Discussions: A Systematic Review	*Canadian Journal of Learning and Technology/La revue canadienne de l'apprentissage et de la technologie*
Journal article	Annand	2019	Limitations of the Community of Inquiry Framework.	*International Journal of E-Learning & Distance Education*
Journal article	Moore, Oliver, & Wang	2019	Setting the pace: examining cognitive processing in MOOC discussion forums with automatic text analysis	*Interactive Learning Environments*,
Journal article	Siburian, Corebima & Saptasari	2019	The Correlation Between Critical and Creative Thinking Skills on Cognitive Learning Results	*Eurasian Journal of Educational Research,*
Journal article	Heilporn, & Lakhal	2020	Investigating the reliability and validity of the community of inquiry framework: An analysis of categories within each presence	*Computers & Education*
Journal article	Foo & Quek		Developing Students' Critical Thinking through Asynchronous Online Discussions: A Literature Review	*Malaysian Online Journal of Educational Technology*
Journal article	Siburian, Corebima, Ibrohim & Saptasari[16]	2019	Analisis Validitas Hasil Pengembangan Perangkat Pembelajaran Ilmu Pengetahuan Lingkungan Berstrategi Inkuiri dan Instrumen Tes Kemampuan Berpikir Kritis, Berpikir Kreatif dan Hasil Belajar Kognitif Mahasiswa	*BIODIK*

[16] Excluded from further analysis due to other language than English.

Thesis	DiPasquale	2017	*Transformative learning and critical thinking in asynchronous online discussions: a systematic review*	Ontario Tech University
Thesis	Moore	2017	*Examining the Influence of Massive Open Online Course Pacing Condition on the Demonstration of Cognitive Presence.*	North Carolina State University
Thesis	Nash	2017	*The Adult Learner's Opinions on the Integration of Critical Thinking Using Technology: A Qualitative Case Study.*	University of Phoenix
Thesis	Butler-Cefalo	2018	*A Case Study of the Application of Community of Inquiry Framework in a Blended Learning Cours*	Northcentral University
Thesis	Melgosa	2018	*Developing a Faculty-Librarian Community of Inquiry: A Blended Learning Approach to Facilitate Information Literacy Education*	University of Calgary
Thesis	Rudolph	2018	*Exploring the Effects of Video Formats on Teaching, Social, and Cognitive Presence in Asynchronous Online Discussions*	The University of Memphis
Thesis	Silva	2018	*Moderating Relationships: Online Learners' Cognitive Presence and Non-designer Instructor's Teaching Presence*	Grand Canyon University
Thesis	Sinning	2018	*Faculty Perceptions of the Transition to Online Instruction: A Case Study.*	University of Phoenix
Thesis	Ervin	2019	*Undergraduate Education Students' Experiences in Online Cooperative Learning Activities: An Embedded Single-Case Study*	Liberty University, Lynchburg
Thesis	Lawson	2019	*Community of Inquiry: Measuring Online Learners' Emotional Presence, Self-Efficacy, and Perceived Quality of Online Learning*	Grand Canyon University
Thesis	Singleton	2019	*Reimagining the Community of Inquiry Model for a Workplace Learning Setting: A Program Evaluation*	University of South Florida
Thesis	Streit	2019	*Online Higher Education: A Correlational Study of Student Engagement, Satisfaction, and Perceptions of Learning*	Grand Canyon University

Part two: Enclosed articles

International Journal of E-Learning & Distance Education
La Revue internationale de l'apprentissage en ligne et de l'enseignement à distance

ISSN: 2292-8588 **Vol. 32, No. 1, 2016**

Critical Thinking in Online Educational Discussions Measured as Progress through Inquiry Phases: A Discussion of the Cognitive Presence Construct in the Community of Inquiry Framework

Jens Breivik

Abstract: The development of critical thinking is a rationale for higher education and an important aspect of online educational discussions. A key component in most accounts of critical thinking is to evaluate the tenability of claims. The *community of inquiry* framework is among the most influential frameworks for research on online educational discussions. In this framework, *cognitive presence* accounts for critical thinking as progress through the following phases of inquiry: triggering event, exploration, integration, and solution

This article discusses the cognitive presence construct as a tool for measuring critical thinking. The article traces the philosophical inspirations of the community of inquiry framework and discusses the construct validity of the cognitive presence construct. Empirical findings enabled by the framework are briefly reviewed and discussed. The author argues that since the cognitive presence construct only to a limited degree addresses the discussants' evaluation of a claim's tenability, the construct possesses weaknesses for assessing critical thinking in discussions. In making this claim, the article contributes to methodological and theoretical discussions about research on critical thinking in online educational discussions.

Keywords: community of Inquiry, cognitive presence, construct validity, e-learning, critical thinking, methodology

Introduction

Online discussions have become a widespread learning activity. The development of critical thinking is a rationale for higher education and a feature frequently examined in research about online discussions. A common method for assessing the quality of online discussions is analysis of discussion transcripts. Nevertheless, critical thinking is not easily defined or operationalized.

Researchers have suggested a large number of different approaches to operationalize critical thinking in online educational discussions. Weltzer-Ward (2011) identified 52 different research frameworks and coding schemes employed between 2002 and 2009 in research on such discussions, although not all of these focused on the critical thinking aspect. According to Weltzer-Ward, a lack of consistent tools hinders comparison of research results and the ability to build on previous analysis. She recommended that researchers should concentrate on a smaller number of frameworks, particularly those that have been most frequently applied in the research field. Among the most frequently applied are frameworks and schemes focusing on inquiry phases (Weltzer-Ward 2011). De Wever, Schellens, Valcke, and Van Keer (2006) discussed 15 different frameworks for content analysis of transcripts from online discussions. According to their research, coherence between the theoretical bases and operationalizations is questionable for a number of the frameworks they examined.

The *Community of Inquiry* (CoI) framework (Garrison, Anderson, & Archer, 1999; Garrison, Anderson, & Archer, 2001; Garrison, 2011) is widely claimed to be a leading research approach

to e-learning in general and online educational discussions in particular (Gašević, Adesope, Joksimović, & Kovanović, 2015; Jézégou, 2010; Shea, 2010; Swan, Garrison, & Richardson, 2009; Weltzer-Ward, 2011). A significant number of studies based on this framework have recently been conducted (Gašević et al., 2015; Horzum & Uyanik, 2015; Lee, 2014; Shea et al., 2014). The framework is also applied through automatic coding software based on learning analytics (Kovanović, Gašević, Joksimović, Hatala, & Adesope, 2015). The CoI framework aims to describe how e-learning can support a collaborative approach to education that promotes deep and meaningful learning (Garrison et al., 1999; Garrison, Anderson, & Archer, 2010). According to the framework, a challenge for online education is to overcome distance and support several kinds of presence. The model suggests the following three distinct but overlapping constructs to assess online educational discussions: *social presence, teaching presence,* and *cognitive presence*.

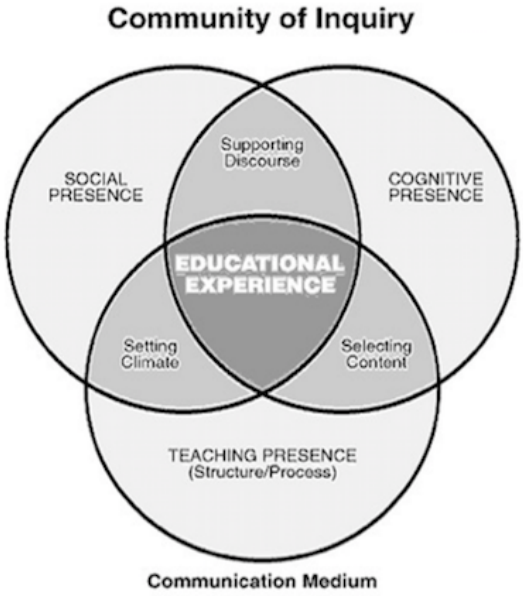


Figure 1: The Community of Inquiry model. Taken from Wikimedia Commons, originally published in Garrison et al. (1999)

Social presence signifies participants' ability to present themselves as "real people" in a purely textual medium and is characterized by emotional expression, open communication, and group cohesion. Teaching presence describes the design and facilitation of the educational experience. Cognitive presence aims to describe higher-order knowledge acquisition and critical thinking as progress through the phases of a triggering event, exploration, integration, and resolution.

The abundance of different frameworks and the variety of different approaches to critical thinking indicate that operationalizing critical thinking is not straightforward. Although the research literature is rich with different approaches, literature that discusses the adequacy and validity of the different frameworks is scarce. For example, there is some discussion about reliability issues (Garrison, Cleveland-Innes, Koole, & Kappelman, 2006), yet little critical discussion of the construct validity of the cognitive presence construct within CoI research. Two articles represent exceptions to this. Rourke and Anderson (2004) noted that one explanation for the observed low levels of cognitive presence might be a result of shortcomings in the coding protocol (p. 11). This observation has not generated much attention. Ho and Swan (2007) suggested an alternative approach to measuring cognitive presence. Their motivation was that established approaches to measuring cognitive presence had been "less successful" or "yielded disappointing results" (p. 4). Inspired by Grice's (1989) cooperative

principle, Ho and Swan (2007) suggested focusing on the communicative quality of discussants' messages rather than on progress toward a solution. Even though their coding scheme has generated interesting empirical findings, the research field has not pursued this approach to measuring cognitive presence. Other researchers focus on how arguments are constructed. Such frameworks (Cho & Jonassen 2002; Weinberger & Fischer, 2006), inspired by Toulmin's (1958/2003) work on informal logic, are commonly used in empirical research (Noroozi, Weinberger, Biemans, Mulder, & Chizari, 2013).

The aim of this article is to discuss the validity of cognitive presence as an operationalization of critical thinking. The research question that guides the discussion is as follows: How adequately is critical thinking in online educational discussions assessed by the cognitive presence construct from the CoI framework?

This examination of the cognitive presence construct contains the following components. First, it traces the philosophical origins of the construct and examines what ideas the framework pursues. Second, the coding scheme is compared with what I will describe as a "minimum conception" of critical thinking. The concept of construct validity guides the comparison. Third, empirical findings based on the framework are briefly reviewed by asking what kind of empirical findings the framework enables, and whether the empirical findings shed light on the measurement tool's capability to assess critical thinking. By discussing these questions, the article contributes to methodological and theoretical debates in the research field. The article does not consider other aspects of the CoI framework, such as the operationalizations of social presence and teaching presence or the idea of distinguishing between different kinds of presences.

Before entering into the discussion of cognitive presence, I will briefly sketch the concepts of construct validity and critical thinking. Construct validity concerns whether an operationalization or coding scheme assesses what it is intended to assess. The development of critical thinking is a key rationale for higher education; however, consensus about how the concept should be interpreted has not been established. In the following, I will discuss some different approaches to critical thinking and sketch what I call a "minimum conception of critical thinking." This minimum conception will be used as a tentative standard against which cognitive presence can be compared.

Background

Construct Validity

Validity in empirical research concerns the quality of the inferences that can be drawn from data (Messick, 1995). This includes considerations of representativeness and possibilities for generalization. When researching abstract and not directly observable phenomena, such as critical thinking, cognitive presence, and deep and meaningful learning, researchers need to operationalize these into observable indicators. Construct validity concerns whether the operationalizations into coding schemes and research tools assess what they are intended to assess. A coding scheme that validly operationalizes a construct will overlap it. The overlapping area represents the features of the construct that the coding scheme validly operationalizes as observable indicators (Kleven, 2008; Messick, 1995).

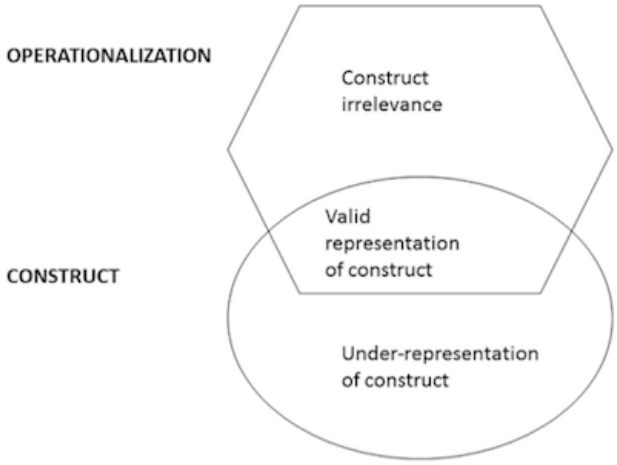


Figure 2: Construct validity

A coding scheme may specify indicators of a construct that are more or less contingent and sufficient to identify a construct. Drawing on Messick (1995) and Kleven (2008), two dimensions, *construct relevance* and *construct representativeness*, constitute the indicators' construct validity.

Construct relevance has to do with whether an operationalization identifies indicators that really belong to the construct. Construct-irrelevant indicators are features that do not represent the construct, although they may be spuriously related to it. If an operationalization of critical thinking identifies indicators that do not necessarily represent critical thinking (e.g., the use of Latin terms), the indicators lack construct relevance. In such cases, utterances that do not demonstrate critical thinking may wrongly be interpreted as showing critical thinking.

Construct representativeness describes whether the operationalization is able to identify necessary, sufficient, or, at least, central features of the construct. When an operationalization fails to identify important features of critical thinking (like precise and unambiguous expressions), the operationalization is considered to be construct under-representative.

A construct-valid operationalization will make it possible to distinguish among and rank empirical instances according to the chosen construct. For research on critical thinking in online educational discussions, discussions (and eventually discussion posts) can thus be ranked according to the levels of critical thinking they contain.

Another aspect of the quality of inferences drawn from data concerns the *reliability of coding*. Reliability of coding describes how consistently different coders interpret and code the same empirical material (inter-rater reliability) and how consistently the same rater codes over time (intra-rater reliability). A number of studies have discussed issues concerning reliability in coding of discussion transcripts (De Wever et al., 2006; Garrison et al., 2006), pointing to the need for appropriate statistical measures. Fewer publications have addressed questions of construct validity, which is the focus of this article.

Critical Thinking

The development of critical thinking is an important rationale for higher education and plays a central role, both as a goal for and as a prerequisite of successful online discussions. A number of adjacent concepts, such as argumentation, reflection, analytical thinking, and rationality, are similar ideals for education. Within the learning sciences, concepts like deep learning (Biggs & Tang, 2007; Marton & Säljö, 1976a, 1976b) and higher-order cognitive processes (Anderson & Krathwohl, 2001) are related educational ideals. According to Siegel (1988/2013) and Lipman

(1991/2003), the idea of critical thinking concerns what good thinking is and is closely related to reasonableness and rationality.

The value and importance of critical thinking and related concepts as educational rationales are supported with several reasons. First, rationality and critical investigation are at the core of science, academic study, and any kind of intellectual activity. Second, rationality is of vital importance in a world where what counts as valid knowledge is rapidly changing. Bluntly formulated, if knowledge keeps no better than fish, a way of checking the durability and validity of knowledge claims is urgently needed. Thus, critical thinking is vital in a knowledge- and innovation-driven economy (for further development of this point, see, for instance, Poce, Corcione, & Iovine, 2012, p. 50). Third, citizens who are critical thinkers and take part in an ongoing rational conversation on societal matters are vital for a democratic society. Dewey (1916/2007) and Habermas (1987) are paradigmatic proponents of this view. Fourth, the ability to think critically and independently is important in order to act as an autonomous subject. According to Siegel (2012) and Bailin and Siegel (2002), to be accepted as an independent center of consciousness, that is, a person, implies being able to perform criteria-governed thinking by oneself about who to be and how best to live. Fifth, providing, analyzing, and justifying claims imply cognitive elaboration and activation of prior knowledge and conceptual schemes, thereby stimulating students' learning (Stegmann, Wecker, Weinberger, & Fischer, 2011).

Despite the widespread acknowledgement of the importance of critical thinking, there seems to be no established consensus regarding how the concept should be understood, much less operationalized. Davies and Barnett (2015), in their introduction to a comprehensive handbook on critical thinking in higher education, put it this way: "After more than four decades of dedicated scholarly work, critical thinking remains more elusive than ever" (p. 3).

A number of different conceptions of critical thinking can be found. By analyzing interviews with academics, Moore (2011a) identified several approaches to how critical thinking is understood. These include *judgment* (evaluating whether claims are sufficiently evidenced); *rationality* (evaluating the validity of inferences drawn from premises and judging whether a conclusion is sufficiently supported by premises); *careful and sensitive readings of texts* (both recognizing how a text's assumptions are supported, and recognizing the text on its own premises and within its own context); *a skeptical and provisional view of knowledge* (which can be described as fallibilism); *self-reflexivity* (the ability to reflect on one's own biases); and *original thinking* (seeing new connections, creativity). Finally, an ethical and activist stance toward contemporary problems is regarded as an additional feature of critical thinking. These themes are addressed in Davies and Barnett's (2015) discussion on whether critical thinking should be regarded as a formal capacity for evaluating arguments; a political and activist orientation focused on identifying missing or concealed dimensions of meaning; or as a concern "with the development of the student as a person" (p. 7) becoming "educated in a modern world" and acting as a "good citizen" (p. 9).

Even though the approaches vary, many accounts of critical thinking point to one core feature as essential: The ability to judge the validity and tenability of a claim, that is, to determine whether a claim is to be considered true. Dewey proposes the following paradigmatic definition of reflective thinking: "active, persistent and careful consideration of any belief or supposed form of knowledge in the light of the grounds that support it and the further conclusions toward which it tends" (Dewey, 1909/1997). This definition was echoed and broadened in a consensus report from the American Philosophical Association: "Critical thinking [is] purposeful, self-regulatory judgment which results in interpretation, analysis, evaluation, and inference, as well as explanation of the evidential, conceptual, methodological, criteriological, or contextual considerations upon which that judgment is based" (Facione, 1990). In line with these definitions, Robert Ennis, a leading scholar in the critical thinking movement, has formulated a brief definition of critical thinking as "reasonable, reflective

thinking focused on deciding what to believe or do" (Ennis, 1991/2015). Ennis builds explicitly on Dewey's work. Ennis' first formulation of what he labeled "a streamlined conception for critical thinking" dates back to 1991. It has been influential in the field and might inform the search for an operational definition when researching online discussions.

The ability to evaluate whether a claim is justified and thus valid will be considered as a minimum conception of critical thinking throughout this article. Evaluating the validity of a claim may include several steps or procedures, including the following:

1. Examine the evidence/grounds that support the claim. Are they tenable, that is, true and valid?
2. Evaluate the relevance of grounds that support the claim. Are the grounds relevant, or do they not support the claim even if they are tenable?
3. Evaluate the inference from the grounds to the claim. Can the claim be deduced with certainty from the grounds, or is the connection merely contingent or probable?
4. Evaluate whether the claim is sufficiently precise and accurately formulated. Does it contain ambiguities that can be interpreted in different ways, thus enabling quasi-agreement or false disagreement?
5. Consider whether a claim and its grounds are biased. Does the claim rely on a proponent's authority or other spurious grounds, rather than on sound backing?
6. Consider the implications of believing in/acting in accordance with the claim. Does the claim lead to objectionable implications?

The list above is my synthesis of arguments from philosophers oriented toward informal logic, such as Dewey (1909/1997), Næss (1953/2005), and Toulmin (1958/2003). The list also echoes philosophical virtues, dating back to the ancient classics, including Plato and Aristoteles.

For the purpose of discussing cognitive presence as a tool to measure critical thinking, further elaboration will be considered excessive. The minimum conception proposes that a core feature of critical thinking is to identify and evaluate whether claims are supported by sufficient and relevant backing. This is in line with Dewey's conception of reflective thinking and certainly makes it relevant in the discussion of the CoI framework, which claims to be consistent with Dewey's account of inquiry and thinking.

Even if one accepts this minimum conception of critical thinking, further vagueness and controversies remain. These include questions about whether critical thinking comprises only skills and abilities or whether attitudes and dispositions toward using such abilities should also be included (Facione, 1990); whether critical thinking can be understood as a generic capacity or whether it is best accounted for as sets of domain-specific capacities (Moore, 2011b); and whether critical thinking should be understood as a formal and epistemic discipline or whether true critical thinking also includes a societal, activist stance (Davies & Barnett, 2015). One source of the fuzziness about how the concept of critical thinking should be understood is that it is investigated from both empirical/descriptive and philosophical/normative angles, and these perspectives have informed each other only to a minor degree (Hyytinen, 2015).

The Community of Inquiry Framework: Cognitive Presence as Phases of Inquiry

Central to the conceptualization of critical thinking in the cognitive presence construct is the practical inquiry model (Garrison, 2011; Garrison et al., 2001; Swan, Garrison, & Richardson, 2009). The model aims toward integrating the following aspects of thinking: perception and reflection, social and private modes of thinking, induction (arrival at generalizations) and deduction (employment of generalizations), intuition and creativity, and action and

deliberation. The practical inquiry model accounts for these aspects as progress through inquiry cycles of the phases included below in Table 1.

Table 1. Cognitive Presence as Phases of Inquiry Process. Based on Garrison (2011) and Shea (2010).

Phase	Indicators
Triggering event	Recognizing the problem Sense of puzzlement
Exploration	Information exchange Suggestions Intuitive leaps
Integration	Connecting ideas Synthesizing ideas Suggesting solutions
Resolution	Applying solutions Testing solutions Defending solutions

These phases form the basis for a coding scheme that enables interpretation and categorization of messages in online discussions that detect which inquiry phase they represent—and thereby the level of critical thinking. All phases of an inquiry process are required in order to achieve high levels of critical thinking. A discussion that reaches a tested solution represents a more advanced level of critical thinking than a discussion that hovers around identifying and understanding a problem. The idea that "deep and meaningful learning does not occur until students move to integration and resolution stages" (Shea et al., 2010, p. 15) is crucial. This way of coding messages makes it possible to classify which level of critical thinking and cognitive presence a discussion as a whole represents. Since all phases of a discussion are essential, there is no option of assessing the quality of a single message. The coding scheme was initially developed in an article by Garrison et al. (1999). A more elaborated and detailed version is found in Shea (2010, p. 20). Similar coding schemes have been developed for the constructs of social presence and teaching presence. A typical research aim has been to investigate relations between cognitive presence and the other constructs.

Arbaugh et al. (2008) have developed a survey instrument based on the CoI framework in order to measure students' conceptions of the different presences in the model and to conduct studies with a multi-institutional scope. The instrument consists of 34 propositions that respondents grade from "strongly disagree" to "strongly agree." Twelve of the questions address cognitive presence and assess how respondents perceive the phases of an inquiry process in their educational experience.[1] Several studies have reported that the instrument is valid in the sense that correlation between the different measured presences occurs as expected (Horzum & Uyanik, 2015; Kozan & Richardson, 2014). In this article, I raise a question about the validity of measures of cognitive presences, even though the measured levels seem to co-variate reasonably with social and teaching presences.

[1] Examples of propositions that describe stages of cognitive presence (Arbaugh et al., 2008) include the following:
"Course activities piqued my curiosity." (Triggering event)
"I felt motivated to explore content-related questions." (Exploration phase)
"Learning activities helped me construct explanations/solutions." (Integration phase)
"I can describe ways to test and apply the knowledge created in this course." (Resolution phase)

Due to the wide reception of the CoI framework, both critiques and suggestions for further developments have been published. In order to demonstrate that cognitive presence as progress through phases remains a stable and defining idea in the CoI framework, the next paragraphs briefly review some of these critique and developments. In a critical review, Rourke and Kanuka (2009) focused on the CoI framework's attempt to promote and research deep and meaningful learning as well as empirical findings showing that integration and solution phases are rarely reached. They argued that the CoI framework has failed as a research strategy by lacking an appropriate and construct-valid tool to measure deep and meaningful learning. Further, they argued that the CoI model fails as a strategy to promote deep and meaningful learning, and they suggested alternative strategies to enhance deep learning based on assessment, reduced content, and confronting misconceptions. Akyol et al. (2009) responded to the critique, partly by dismissing the methodology of the review and partly by insisting that the CoI framework focuses on learning *processes*, not learning *results*. Jézégou (2010) criticized the CoI framework for weak and unclear theoretical foundations. Garrison (2013) responded by explicating constructivism as the theoretical background for the CoI framework and announcing metacognition as a supplemental construct in the model. Xin (2012) suggested that many utterances in a discussion contain social, teaching, and cognitive functions at the same time. This challenges the idea of separating such presences.

Nevertheless, few critics of the CoI framework have questioned whether cognitive presence and thereby critical thinking is adequately operationalized and assessed as progress through the phases of inquiry.

Recently, supplements to the CoI model have been suggested. As mentioned above, one is the construct of *metacognition*, which focuses on one's own and peers' knowledge, monitoring, and reflection of cognition, and claims that self-corrective strategies are integral to inquiry (Akyol & Garrison, 2011; Garrison & Akyol, 2015). Another recent suggestion is to supplement the model with the construct of *learning presence* (Shea & Bidjerano, 2010; Shea et al., 2012; Shea et al., 2014). This construct emphasizes the value of self- and co-regulation and monitoring of learning. Although there are clear similarities between these suggested new constructs, there has been some controversy among their proponents. Both groups assert that the goal of the educational process is cognitive presence, understood and measured as progress through the phases of inquiry. Redmond (2014) suggested a similar idea by claiming that reflection (understood as self- and co-reflection and awareness along with regulation of cognition) is definitive for cognitive presence and can be included in the established coding scheme. Despite these critiques and developments, cognitive presence operationalized as progress through phases has remained a stable element of the CoI framework.

How are Ideas from the Philosophical Sources Pursued in the CoI Framework?

The CoI framework has several theoretical foundations. The main inspiration for the CoI framework is Dewey's thoughts on education and thinking and the relation between them (Garrison et al., 1999; Garrison, 2011; Swan et al., 2009). One of Dewey's central motives (Dewey, 1909/1997, 1920/1971, 1933/1986) was a rejection of what he called "false dualisms." Traditional epistemologies tend to set up *world* and *mind*, *thinking* and *action*, and the *individual* and the *social* as contrasts that shape the way we conceptualize the world. Dewey aimed at a naturalistic account of thinking and knowing by putting knowledge and the knowledge-seeking subject back into their surroundings (Rorty, 1986). Thinking should not be understood as an inner process within a solitary subject disengaged from the world, but as toil to overcome and cope with problems in the world. In this view, thinking is closely related to problem solving. According to this view, an inquiry process is typically initiated when a problem occurs as a situation where the actor's established responses to the environment are inadequate. Latter phases in the process include elaboration and testing of hypothetical solutions. Dewey took the scientific method, and especially Darwin's contributions to biology, as ideals for investigation, but he pointed out that there are no essential differences between

everyday reasoning and scientific reasoning. What characterizes scientific reasoning is a higher degree of precision and abstraction. For all kinds of reasoning, conclusions rely on testing in the real world rather than on speculative evaluation. Further, the development of reflection and reasoning is an inherent goal for education, which is more crucial than teaching students to memorize facts.

Lipman (Lipman, 1991/2003; Lipman, Sharp, & Oscanyan, 1980) introduced "community of inquiry" as an educational concept, emphasizing that dialogues are important for education. Originally, the concept was introduced by Peirce, a pragmatist philosopher and predecessor of Dewey. According to Peirce, the concept illustrates that scientific inquiry normally is conducted in collaboration with colleagues and peers. Criticizing and defending assumptions and hypotheses in a community of peers is thus the primordial mode of inquiry. Inspired by Dewey, Lipman (1991/2003) developed the concept of community of inquiry into an educational concept. According to Lipman, education should stimulate students' ability to reflect. Furthermore, Lipman acknowledged that dialogues performed in a community of peers strengthen students' reasoning and critical thinking abilities. In line with this, Lipman, Sharp, and Oscanyan (1980) have developed a program for developing pupils' reflective reasoning by dialogical means, which is labeled Philosophy for Children (sometimes abbreviated as P4C). The concept of community of inquiry is commonly associated with the movement for promoting philosophy with children.

Another tradition that has inspired the CoI framework focuses on the concept of deep learning, as opposed to surface learning (Biggs & Tang, 2007; Marton & Säljö, 1976a, 1976b). The concept of deep learning connects cognitive presence with educational rationales, such as higher-order thinking and critical thinking.

Other than Lipman's influence, references to the general literature on critical thinking in education are scarce in the research literature on the CoI framework. Ideas from researchers like Ennis (1989), Siegel (1988/2013), and Toulmin (1958/2003) might have contributed to a broader conception of critical thinking.

The element from Dewey pursued in the cognitive presence coding scheme is the idea that inquiry proceeds through defined phases. In Dewey's account, progress in inquiry is used to illustrate how thinking (i.e., problem solving and inquiry) is embedded in practical contexts and is a response to obstacles to actors' enterprises in their surrounding.[2] This description of how thinking, and thus inquiry, emerges as a response to obstacles in the surroundings is taken as an account of the quality of the process. One might ask if this represents confusion between a description of how thinking is initiated and an evaluation of how advanced or successful the thinking is.

Proponents of the CoI model emphasize that Dewey's insights show the importance of dialogue and community to thinking, inquiry, and education. However, the cognitive presence construct does not necessarily highlight the social aspect of thinking. One might picture an individual in solitude progressing through the phases of inquiry. Progress through phases of inquiry as an operationalization of critical thinking and interpretation of Dewey's philosophy does not emphasize the social and dialogical aspect of thinking.

---

[2] The description of phases is built on a description of such phases of inquiry described in the 1909 edition of Dewey's book *How We Think*: 1) occurrence of a difficulty; 2) definition of the difficulty; 3) suggestion of possible solutions; 4) rational elaboration of bearings; and 5) corroboration and founding of concluding belief. However, in the 1933 edition of the book, Dewey makes it clear that these phases do not necessarily proceed in a fixed order, but should rather be regarded as aspects or modes of thinking and reflection as problem solving. The points Dewey emphasizes is that problems are intertwined with the situations they occur in, and that reflection and problem solving are best understood as a process of suggesting and testing solutions to such problems. It may be a bit awkward that Garrison and colleagues' references to *How We Think* are given for the revised edition from 1933 when the points made seem to be more in line with the 1909 edition.

Is the Cognitive Presence Coding Scheme Construct Valid as an Operationalization of Critical Thinking?

Construct validity concerns whether an operationalization or coding scheme measures what it intends to measure. Cognitive presence in the CoI framework is meant to be closely related to critical thinking. As previously described, a minimum conception of critical thinking highlights checking of claims tenability, and the cognitive presence construct operationalizes critical thinking as progress through phases of inquiry. How does progress through phases resemble checking the tenability of claims?

One answer to this question is *not very well!* Progress from the triggering event via exploration toward the integration may take place without any investigation of how well relevant claims and arguments are supported. The indicators that are associated with these phases, like "sense of puzzlement," "information exchange," and "connecting ideas" (see Table 1), do not necessarily imply any checking of the claims' tenability. Indicators used to describe the solution phase, like "testing solutions" and "defending solutions," may imply the evaluation of tenability.

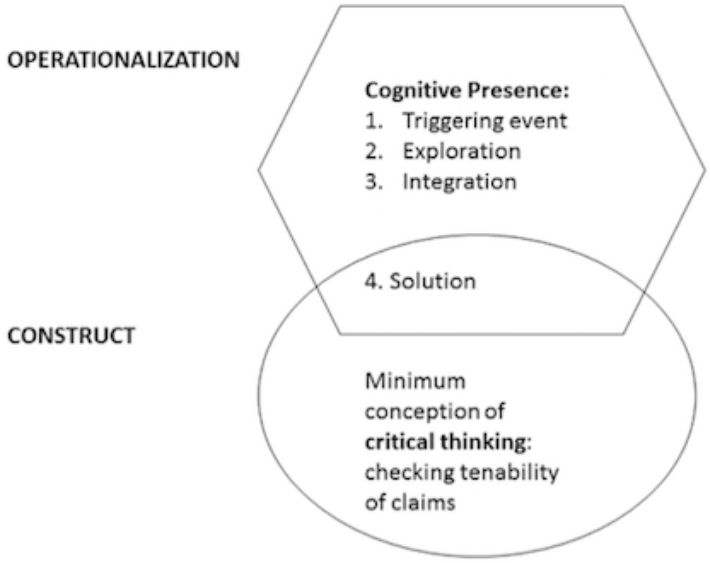


Figure 3: Construct validity of cognitive presence coding scheme as a measure of critical thinking. Concordance between checking tenability of claims and progress through phases of inquiry.

Figure 3 illustrates that there is limited necessary overlap between the cognitive presence coding scheme and the minimum conception of critical thinking. Critical thinking may take place independently of progress through predefined phases as a necessary characteristic, indicating that the operationalization underrepresents what it aims to measure. Similarly, problem solving through predefined phases of inquiry may take place without any critical thinking necessarily taking place, thus indicating the irrelevance of the operationalization. It is important to note the term *necessary* in the description above. It may happen that critical thinking actually takes place during progress through inquiry phases. However, such occasional connection between critical thinking and progress through inquiry phases is not necessary or sufficient to describe attributes of either critical thinking or progress through inquiry phases. This indicates that the coding scheme is both construct irrelevant and construct under-representative, which implies weak construct validity for the cognitive coding scheme.

Critical thinking may be defined by other characteristics than those suggested in my minimum conception, i.e., evaluation of claims' tenability. Alternative characteristics of critical thinking

might be precision, fallibilism, perspective awareness, intellectual integrity, un-tendentiousness, or ability to draw logical valid inferences. Defining critical thinking by such virtues will lead to similar conclusions regarding the construct validity of progress through phases as a measure of critical thinking. Progress toward a solution as an indicator for the level of critical thinking implies that thoroughness and meticulousness are less desired. Focusing on reaching a solution may imply that the framework emphasizes aspects of a discussion that do not necessarily represent critical thinking.

A construct-valid coding scheme for assessing critical thinking should have the capacity to distinguish discussions with high levels of critical thinking from discussions with low levels of critical thinking. It is questionable whether the coding scheme based on cognitive presence has this capacity. Discussions that do not reach a solution but that concern complex problems would be rated as showing lower levels of critical thinking than discussions in which (even wrong) solutions are reached to less demanding questions. Overvaluing solutions to intellectual problems might favor a less thoughtful approach or reward solutions even when they rely on poor understanding, weak evidence, unclear concepts, and so forth.

Results from Empirical Studies

As previously mentioned, a common finding in reviews of CoI research is that online educational discussions rarely reach phases of integration and resolution, but instead tend to hover around the initial phases of the inquiry process (Garrison et al., 2001; Garrison, 2011; Kanuka & Anderson, 1998; Rourke & Kanuka, 2009; Zydney, deNoyelles, & Kyeong-Ju Seo, 2012). In keeping with the idea that deep and meaningful learning takes place when students reach the integration and resolution phases, several explanations might be possible (Garrison et al., 2001; Rourke & Anderson, 2004).

One way of explaining low incidences of discussion posts in the integration and resolution phases is to presume that online discussions fail as a strategy to promote cognitive presence and thereby critical thinking (Rourke & Kanuka, 2007, 2009). Another way of explaining the measured low levels of cognitive presence is to interpret it as a result of unsatisfactory levels of teaching presence, and thereby a result of weak facilitation. Discussion tasks designed with clear expectations about outcomes and facilitation that guides discussion forward, might drive discussions toward the resolution phase (Garrison, 2011; Swan et al., 2009).

Yet another strategy is to expect higher levels of cognitive presence to occur in media and learning activities other than online discussion:

> These more abstract phases of knowledge construction will not be most evident in student interactive discourse (threaded discussions) but should, instead, be evident in activities designed to allow for their demonstration, such as integrative papers, projects, case studies and the like (Shea, 2010, Section 4.4).

According to this view, online discussions do not fulfill their intended purpose; however, low observed levels of cognitive presence are explained within the CoI framework. The previous discussion of the construct validity of the cognitive presence measure suggests that low observed levels of cognitive presence might also be a result of the tool's weak ability to measure critical thinking. This implies that it is impossible to know whether the low observed levels of cognitive presence could be explained as weaknesses in online discussions as an arena that fosters critical thinking, weaknesses in facilitation, or as weaknesses in the measurement tool (Rourke & Anderson, 2004).

Discussion and Implications

The above discussion has addressed how the cognitive presence construct measures critical thinking. To sum up, first, the cognitive presence measurement tool takes Dewey's description of how thinking and inquiry are initiated as responses to obstacles in the surroundings—following certain steps of problem solving—as an account of what critical thinking is. This way

of reading Dewey does not distinguish between how thinking, inquiry, and problem solving are initiated and what good critical thinking is. An alternative approach, based on Dewey (1909/1997) and pursued by Ennis (1991/2015), is to consider critical thinking as the evaluation of a claim's tenability.

Second, the cognitive presence measurement tool does not target discussants' evaluation of a claim's tenability as the key feature of critical thinking. Compared to a minimum conception of critical thinking that takes "deciding what to believe" as a hallmark, the coding scheme has weak construct validity, and the operationalized indicators—progress through phases of inquiry—might be considered both irrelevant and unrepresentative.

Third, empirical studies based on a cognitive presence measurement tool show that high levels of cognitive presence are rarely found in online educational discussions. Rather than initiating investigations about the validity of the coding schemes, such findings are explained by ad hoc hypotheses, such as explaining low observed levels of cognitive presence as a result of weak facilitation or expecting that students will express higher levels of cognitive presence in other learning activities.

Discussion

The CoI framework and the cognitive presence coding scheme are frequently applied in empirical research (Gašević, Adesope, Joksimović, & Kovanović, 2015; Weltzer-Ward, 2011). Do the wide acceptance of the framework and the operationalization of cognitive presence indicate that the argument presented in this article is erroneous? At least two objections regarding the soundness of this article's claims may be raised. First, is it really critical thinking that the CoI framework aims to measure with the cognitive presence construct? Second, even if the cognitive presence construct does not describe critical thinking (as defined here), cognitive presence might define and measure other important educational rationales.

In the initial and seminal articles by Garrison et al. (1999, 2001), it is made quite clear that cognitive presence is proposed as an operationalization of critical thinking: "cognitive presence as assessed and defined by the phases [represent] a generalized model of critical thinking" (Garrison et al., 2001, p. 8). Moreover, these articles discussed established characteristics of critical thinking (accuracy, logic, etc.) and rejected these as being too "algorithmic" to be a suitable strategy to operationalize critical thinking (Garrison et al., 2001, p. 12). Instead, the founders of the CoI framework wanted to focus on discussants that "demand reason for beliefs [… in a] self-judging community" (Garrison et al., 2001, p. 12). The claim in this article is that such a conception of critical thinking is not measured by the cognitive presence indicators in a manner that upholds construct validity. Such an approach is described by what I have called a minimum conception of critical thinking. Operationalizations of discussants' demands of reasons for beliefs—based on informal logic (Toulmin, 1958/2003), as opposed to "algorithmic" approaches to cognition—are proposed in the research field in several frameworks (Clark, Sampson, Weinberger, & Erkens, 2007; Noroozi, Weinberger, Biemans, Mulder, & Chizari, 2012; Weinberger & Fischer, 2006) and may represent an alternative approach to measuring cognitive presence.

What other kinds of educational rationales do the cognitive presence indicators possibly measure, if not critical thinking? Problem solving is commonly associated and shares a number of overarching themes with critical thinking (McCormick, Clark, & Raines, 2015). The CoI literature mentions deep and meaningful learning (Biggs & Tang, 2007) as an associated educational rationale. Higher-order knowledge acquisition is also a candidate (Anderson & Krathwohl, 2001). One might also expect that high levels of cognitive presence coincide with student engagement (Kahu, 2013; Kuh, 2009), cognitive engagement, and students spending time on task rather than drifting into distractions. Nevertheless, a number of taxonomies and tools have been developed to measure such educational rationales and it has not been

established that cognitive presence measured as progress through phases of inquiry represents a construct-valid measurement for any of them.

Implications

The argument in this article is addressed primarily toward the research community. If the argument is found persuasive, it may spark some debate in parts of the community of inquirers on collaboration and critical thinking in digital educational settings. The argument may be refuted with an adequate explication of the validity of cognitive presence, or it may stimulate further development of the CoI framework.

Other possible implications concern how to guide and facilitate educational discussions. If critical thinking is more profoundly accounted for as checking a claim's tenability than as progression toward a solution, this might influence instructions and facilitation of discussions.

References

Akyol, Z., Arbaugh, J. B., Cleveland-Innes, M., Garrison, D. R., Ice, P., Richardson, J. C., & Swan, K. (2009). A Response to the Review of the Community of Inquiry Framework. *Journal of Distance Education*, *23*(2), 123-135.

Akyol, Z., & Garrison, D. R. (2011). Assessing Metacognition in an Online Community of Inquiry. *The Internet and Higher Education*, *14*(3), 183-190. doi:http://dx.doi.org/10.1016/j.iheduc.2011.01.005

Anderson, L. W., & Krathwohl, D. R. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. New York: Longman.

Arbaugh, J. B., Cleveland-Innes, M., Diaz, S. R., Garrison, D. R., Ice, P., Richardson, J. C., & Swan, K. P. (2008). Developing a Community of Inquiry Instrument: Testing a Measure of the Community of Inquiry Framework Using a Multi-Institutional Sample. *The Internet and Higher Education*, *11*(3-4), 133-136. doi:http://dx.doi.org/10.1016/j.iheduc.2008.06.003

Bailin, S., & Siegel, H. (2002). Critical Thinking. In N. Blake, P. Smyers, R. Smith, & P. Standish (Eds.), *The Blackwell guide to the philosophy of education*. Blackwell Publishing.

Biggs, J. B., & Tang, C. S.-k. (2007). *Teaching for quality learning at university: What the student does*. Maidenhead: McGraw-Hill/Society for Research into Higher Education & Open University Press.

Cho, K.-L., & Jonassen, D. H. (2002). The Effects of Argumentation Scaffolds on Argumentation and Problem Solving. *Educational Technology Research and Development*, *50*(3), 5-22. doi:10.1007/bf02505022

Clark, D., Sampson, V., Weinberger, A., & Erkens, G. (2007). Analytic Frameworks for Assessing Dialogic Argumentation in Online Learning Environments. *Educational Psychology Review*, *19*(3), 343-374. doi:10.1007/s10648-007-9050-7

Davies, M., & Barnett, R. (2015). Introduction. In M. Davies & R. Barnett (Eds.), *The Palgrave handbook of critical thinking in higher education*.

De Wever, B., Schellens, T., Valcke, M., & Van Keer, H. (2006). Content Analysis Schemes to Analyze Transcripts of Online Asynchronous Discussion Groups: A Review. *Computers & Education*, *46*(1), 6-28. doi: http://dx.doi.org/10.1016/j.compedu.2005.04.005

Dewey, J. (1909/1997). *How we think*. Mineola, N.Y.: Dover Publ.

Dewey, J. (1920/1971). *Reconstruction in philosophy*. Boston: Beacon Press.

Dewey, J. (1933/1986). How we think (rev. ed.) *The later works: 1925–1953: Vol. 8: 1933*. Carbondale, Ill: Southern Illinois University Press.

Dewey, J. (2007). *Democracy and education: An introduction to the philosophy of education*. [Sioux Falls]: NuVision.

Ennis, R. H. (1989). Critical Thinking and Subject Specificity: Clarification and Needed Research. *Educational researcher*, *18*(3), 4–10.

Ennis, R. H. (1991/2015). Critical thinking: A Streamlined Conception. In M. Davies & R. Barnett (Eds.), *The Palgrave handbook of critical thinking in higher education* (pp. 31-47). New York: Palgrave Macmillan US.

Facione, P. A. (1990). Critical Thinking: A Statement of Expert Consensus for Purposes of Educational Assessment and Instruction. *Research Findings and Recommendations*.

Garrison, Anderson, T., & Archer, W. (1999). Critical Inquiry in a Text-Based Environment: Computer Conferencing in Higher Education. *The Internet and Higher Education*, *2*(2-3), 87-105. doi:10.1016/s1096-7516(00)00016-6

Garrison, Anderson, T., & Archer, W. (2001). Critical Thinking, Cognitive Presence, and Computer Conferencing in Distance Education. *American Journal of Distance Education*, *15*(1), 7-23.

Garrison, D. R. (2007). Online Community of Inquiry Review: Social, Cognitive, and Teaching Presence Issues. *Journal of Asynchronous Learning Networks*, *11*(1), 61-72.

Garrison, D. R. (2011). *E-learning in the 21st century*. New York: Routledge.

Garrison, D. R. (2013). Theoretical Foundations and Epistemological Insights of the Community of Inquiry. In Z. Akyol & D. R. Garrison (Eds.), *Educational communities of inquiry theoretical framework, research and practice*. Hershey: Information scinece Reefernce, IGI Global.

Garrison, D. R., & Akyol, Z. (2015). Toward the Development of a Metacognition Construct for Communities of Inquiry. *The Internet and Higher Education*, *24*, 66-71. doi:http://dx.doi.org/10.1016/j.iheduc.2014.10.001

Garrison, D. R., Anderson, T., & Archer, W. (2010). The First Decade of the Community of Inquiry Framework: A Retrospective. *The Internet and Higher Education*, *13*(1), 5-9.

Garrison, D. R., Cleveland-Innes, M., Koole, M., & Kappelman, J. (2006). Revisiting methodological issues in transcript analysis: Negotiated coding and reliability. *The Internet and Higher Education*, *9*(1), 1–8. doi:http://dx.doi.org/10.1016/j.iheduc.2005.11.001

Gašević, D., Adesope, O., Joksimović, S., & Kovanović, V. (2015). Externally Facilitated Regulation Scaffolding and Role Assignment to Develop Cognitive Presence in Asynchronous Online Discussions. *The Internet and Higher Education*, *24*, 53-65. doi: http://dx.doi.org/10.1016/j.iheduc.2014.09.006

Grice, P. (1989). *Studies in the way of words*. Cambridge, Mass.: Harvard University Press.

Habermas, J., & Habermas, J. (1987). *The theory of communicative action: 2: Lifeworld and system: A critique of functionalist reason*. London: Heinemann.

Ho, C.-H., & Swan, K. (2007). Evaluating Online Conversation in an Asynchronous Learning Environment: An Application of Grice's Cooperative Principle. *The Internet and Higher Education*, *10*(1), 3–14. doi: http://dx.doi.org/10.1016/j.iheduc.2006.11.002

Horzum, M. B., & Uyanik, G. K. (2015). An Item Response Theory Analysis of the Community of Inquiry Scale. *The International Review of Research in Open and Distributed Learning*, *16*(2).

Hyytinen, H. (2015). *Looking beyond the obvious: Theoretical, empirical and methodological insights into critical thinking*. PhD thesis delivered to University of Helsinki

Jézégou, A. (2010). Community Of Inquiry in E-Learning: A Critical Analysis of the Garrison and Anderson Model. *Journal of Distance Education*, *24*(3), 18–27.

Kahu, E. R. (2013). Framing Student Engagement in Higher Education. *Studies in Higher Education*, *38*(5), 758–773. doi: 10.1080/03075079.2011.598505

Kanuka, H., & Anderson, T. (1998). Online Social Interchange, Discord, and Knowledge Construction. *Journal of Distance Education*, *13*(1).

Kleven, T. A. (2008). Validity and Validation In Qualitative and Quantitative Research. *Nordic Studies in Education*, *28*(03).

Kovanović, V., Gašević, D., Joksimović, S., Hatala, M., & Adesope, O. (2015). Analytics of Communities of Inquiry: Effects of Learning Technology Use on Cognitive Presence in Asynchronous Online Discussions. *The Internet and Higher Education*, *27*, 74-89. doi: http://dx.doi.org/10.1016/j.iheduc.2015.06.002

Kozan, K., & Richardson, J. C. (2014). New Exploratory and Confirmatory Factor Analysis Insights into the Community of Inquiry Survey. *The Internet and Higher Education*, *23*, 39-47. doi: http://dx.doi.org/10.1016/j.iheduc.2014.06.002

Kuh, G. D. (2009). What Student Affairs Professionals Need to Know about Student Engagement. *Journal of College Student Development*, *50*(6), 683–706. doi: 10.1353/csd.0.0099

Lee, S.-M. (2014). The Relationships Between Higher Order Thinking Skills, Cognitive Density, and Social Presence in Online Learning. *The Internet and Higher Education*, *21*, 41–52. doi: http://dx.doi.org/10.1016/j.iheduc.2013.12.002

Lipman, M. (1991/2003). *Thinking in education* (2nd ed.). Cambridge: Cambridge University Press.

Lipman, M., Sharp, A. M., & Oscanyan, F. S. (1980). *Philosophy in the classroom*. Philadelphia: Temple University Press.

Marton, F., & Säljö, R. (1976a). On Qualitative Differences in Learning: I: Outcome and Process. *British Journal of Educational Psychology*, *46*(1), 4–11. doi: 10.1111/j.2044-8279.1976.tb02980.x

Marton, F., & Säljö, R. (1976b). On Qualtitative Differences in Learning – II: Outcome as a Function of the Learner's Conception of the Task. *British Journal of Educational Psychology*, *46*(2), 115-127. doi: 10.1111/j.2044-8279.1976.tb02304.x

McCormick, N. J., Clark, L. M., & Raines, J. M. (2015). Engaging Students in Critical Thinking and Problem Solving: A Brief Review of the Literature. *Journal of Studies in Education*, *5*(4), 100-113.

Messick, S. (1995). Validity of Psychological Assessment: Validation Of Inferences from Persons' Responses and Performances as Scientific Inquiry into Score Meaning. *American Psychologist*, *50*(9), 741-749. doi: 10.1037/0003-066X.50.9.741

Moore, T. (2011). Critical thinking: Seven definitions in Search of a Concept. *Studies in Higher Education*, *38*(4), 506–522. doi: 10.1080/03075079.2011.586995

Moore, T. J. (2011). Critical thinking and Disciplinary Thinking: A Continuing Debate. *Higher Education Research & Development, 30*(3), 261-274. doi: 10.1080/07294360.2010.501328

Noroozi, O., Weinberger, A., Biemans, H. J. A., Mulder, M., & Chizari, M. (2012). Argumentation-Based Computer Supported Collaborative Learning (ABCSCL): A Synthesis of 15 Years of Research. *Educational Research Review*, *7*(2), 79-106. doi: http://dx.doi.org/10.1016/j.edurev.2011.11.006

Noroozi, O., Weinberger, A., Biemans, H. J. A., Mulder, M., & Chizari, M. (2013). Facilitating Argumentative Knowledge Construction Through a Transactive Discussion Script in CSCL. *Computers & Education*, *61*(0), 59–76. doi: http://dx.doi.org/10.1016/j.compedu.2012.08.013

Næss, A. (1953/2005). *The selected works of Arne Naess: Vol. 1: Interpretation and preciseness: A contribution to the theory of communication*. Dordrecht: Springer.

Poce, A., Corcione, L., & Iovine, A. (2012). Content Analysis and Critical Thinking. An Assessment Study. *Cadmo*.

Redmond, P. (2014). Reflection as an Indicator of Cognitive Presence. *E-Learning and Digital Media*, *11*(1), 46-58.

Rorty, R. (1986). Introduction. In *Dewey: The later works: 1925–1953: Vol. 8: 1933*. Carbondale, Ill: Southern Illinois University Press.

Rourke, L., & Anderson, T. (2004). Validity in Quantitative Content Analysis. *Educational Technology Research and Development*, *52*(1), 5-18. doi: 10.1007/bf02504769

Rourke, L., & Kanuka, H. (2007). Barriers to Online Critical Discourse. *International Journal of Computer-Supported Collaborative Learning*, *2*(1), 105–126. doi: 10.1007/s11412-007-9007-3

Rourke, L., & Kanuka, H. (2009). Learning in Communities of Inquiry: A Review of the Literature. *Journal of Distance Education*, *23*(1), 19-48.

Shea, P. (2010). A Re-Examination of the Community of Inquiry Framework: Social Network and Content Analysis. *The Internet and Higher Education*, *13*(1), 10.

Shea, P., & Bidjerano, T. (2010). Learning Presence: Towards a Theory of Self-Efficacy, Self-Regulation, and the Development of a Communities of Inquiry in Online and Blended Learning Environments. *Computers & Education*, *55*(4), 1721–1731. doi: http://dx.doi.org/10.1016/j.compedu.2010.07.017

Shea, P., Hayes, S., Smith, S. U., Vickers, J., Bidjerano, T., Pickett, A., . . . Jian, S. (2012). Learning Presence: Additional Research on a New Conceptual Element Within the Community of Inquiry (Coi) Framework. *The Internet and Higher Education*, *15*(2), 89–95. doi: http://dx.doi.org/10.1016/j.iheduc.2011.08.002

Shea, P., Hayes, S., Uzuner-Smith, S., Gozza-Cohen, M., Vickers, J., & Bidjerano, T. (2014). Reconceptualizing the Community of Inquiry Framework: An Exploratory Analysis. *The Internet and Higher Education, 23*(0), 9-17. doi: http://dx.doi.org/10.1016/j.iheduc.2014.05.002

Shea, P., Hayes, S., Vickers, J., Gozza-Cohen, M., Uzuner, S., Mehta, R., . . . Rangan, P. (2010). A Re-Examination of the Community of Inquiry Framework: Social Network and Content Analysis. *The Internet and Higher Education, 13*(1–2), 10–21. doi: http://dx.doi.org/10.1016/j.iheduc.2009.11.002

Siegel, H. (1988/2013). *Educating reason*. Routledge.

Siegel, H. (2012). Education as Initiation into the Space of Reasons. *Theory and Research in Education*, *10*(2), 191–202. doi: 10.1177/1477878512446542

Stegmann, K., Wecker, C., Weinberger, A., & Fischer, F. (2011). Collaborative Argumentation and Cognitive Elaboration in a Computer-Supported Collaborative Learning Environment. *Instructional Science*, *40*(2), 297–323. doi: 10.1007/s11251-011-9174-5

Swan, K., Garrison, D., & Richardson, J. (2009). A Constructivist Approach to Online Learning: The Community of Inquiry Framework. *Information technology and constructivism in higher education: Progressive learning frameworks.* Hershey, PA: IGI Global.

Toulmin, S. (1958/2003). *The uses of argument*. Cambridge: Cambridge University Press.

Weinberger, A., & Fischer, F. (2006). A Framework to Analyze Argumentative Knowledge Construction in Computer-Supported Collaborative Learning. *Computers & Education*, *46*(1), 71–95. doi: http://dx.doi.org/10.1016/j.compedu.2005.04.003

Weltzer-Ward, L. (2011). Content Analysis Coding Schemes for Online Asynchronous Discussion. *Campus-Wide Information Systems*, *28*(1), 56-74.

Xin, C. (2012). A Critique of the Community of Inquiry Framework. *International Journal of E-Learning & Distance Education, 26*(1).

Zydney, J. M., deNoyelles, A., & Kyeong-Ju Seo, K. (2012). Creating a Community of Inquiry in Online Environments: An Exploratory Study on the Effect of a Protocol on Interactions Within Asynchronous Discussions. *Computers & Education*, *58*(1), 77-87. doi: http://dx.doi.org/10.1016/j.compedu.2011.07.009

Author

Jens Breivik, Center for Teaching, Learning and Technology, UiT - The Arctic University of Norway. E-mail: jens.breivik@gmail.com

Nordic Journal of Digital Literacy

UNIVERSITETSFORLAGET

RESEARCH PUBLICATION

Vol. 15, No. 1-2020, p. 8–23
ISSN Online: 1891-943X
https://doi.org/10.18261/issn.1891-943x-2020-01-02

Argumentative patterns in students' online discussions in an introductory philosophy course

Micro- and macrostructures of argumentation as analytic tools

Jens Breivik
Associate Professor, Department of Education, UiT – The Arctic University of Norway
jens.breivik@uit.no

Abstract

Online discussions are commonly used as learning activities in higher education. One of the rationales behind their use is to enhance students' competence in critical thinking and rational argumentation. In the research field, several approaches to critical thinking and rational argumentation are suggested, and several frameworks for analyzing online educational discussions are employed. In this article, online discussions from an introductory philosophy course are analyzed. The microstructure of arguments (how arguments are backed) and the macrostructure of argumentation (how arguments are linked together in chains of arguments and counterarguments) are used as analytic tools. The categories for analysis are based on Toulmin's argument model. The aim here is twofold. First, the article explores what occurs in online discussions in an introductory philosophy course where competence in argumentation is a specific learning objective, analyzed using the categories of the microstructure of arguments and the macrostructure of argumentation. Second, the article discusses how suitable the categories from Toulmin's model are for such analysis. The analysis reveals that the students eagerly discussed the topic, showed an understanding of the topic, and employed subject knowledge. Yet, their discussion posts tended to be associative and unaddressed. The categories of the microstructure of arguments and the macrostructure of argumentation proved powerful tools for analysis. The analysis coincides with the students' and teachers' own evaluation of argumentation in the discussions, yet it provides a more justified, detailed picture of the strengths and weaknesses in the students' argumentation. Nevertheless, important qualities of the discussion are not revealed by these categories. One recommendation for teaching and facilitation is to provide students with an elaborated conception of rational argumentation.

Keywords
Argumentation, Online discussions, Arguments' micro-structure, Argumentation's macro-structure, Toulmin-model

Introduction

Online discussion is a common learning activity in higher education (Wise & Paulus, 2016). Typically, students discuss topics asynchronously on discussion boards set up in learning management systems (LMSs) or other platforms. In this study, discussions were set up as a mandatory learning activity in an introductory philosophy course to enhance students' competence in argumentation.

How did the students construct and place their arguments in the online discussions? Moreover, what are suitable categories to analyze argumentation in online discussions? In this article, these questions guide the analysis of discussion transcripts using the categories

of the microstructure of arguments and the macrostructure of argumentation. The first question derives from an urge to understand how online discussions contribute to students' learning and a desire to better facilitate online discussions for enhanced learning. The second question derives from theoretical and methodological discussions in the research field about how to analyze argumentation in educational online discussions.

In education, online discussion serves different purposes—from socialization to the dissemination of information to the development of new knowledge (Salmon, 2013). Developing new knowledge, scrutinizing established conceptions, and contrasting conflicting views all require the ability to reflect and think critically, propose sound arguments, and evaluate the tenability of arguments (Andriessen & Baker, 2014). Deep and higher-order learning require the ability to provide reasons for, evaluate, and justify knowledge claims and not merely the ability to memorize and reproduce knowledge (Biggs & Tang, 2007; Marton & Säljö, 1976). The quality of arguments students are able to articulate reflect the quality of their learning (Andriessen & Baker, 2014). Despite the wide acceptance of critical thinking and rational argumentation as learning objectives, research has shown students do not develop these competencies as desired (Arum & Roksa, 2011; Cahill & Bloch-Schulman, 2012).

There are several approaches to critical thinking and rational argumentation in the research field; consequently, there are several approaches to teaching and assessing rational argumentation and critical thinking. Most studies on critical thinking maintain that the ability to justify, evaluate, and provide reasons for knowledge claims is central (Dewey, 1998; Ennis, 1989; Facione & Facione, 2007; Siegel, 1988/2013; Toulmin, 1958/2003). Therefore, competence in critical thinking and rational argumentation provides students with powerful strategies for independent thinking and deep learning.

The development of students' critical thinking and their ability to argue rationally are vital for higher education, and are commonly stated learning objectives. Participation in discussions provides an opportunity to practice rational argumentation, and to elaborate and contest discussants' views. Online discussions—written and asynchronous—may enable in-depth debate and thoughtful learning. Educators and researchers strive to explore how participation in online written discussions influences students' learning and understanding of subject matter (Kovanović, Gašević, Joksimović, Hatala, & Adesope, 2015; Loncar, Barrett, & Liu, 2014; Schindler & Burkholder, 2014; Tsai & Tsai, 2014; Wise & Paulus, 2016; Zheng & Warschauer, 2015). Expectations about how online discussions may support learning have been high; however, the research includes mixed documentation of the significance of online discussions vis-à-vis learning (Rourke & Kanuka, 2009; Thomas, 2013). Previous research has shown there are challenges with online discussion, such as low activity, students not responding to each other's posts, and surface-level rather than in-depth discussion. Understanding what occurs in the discussions may inform and thus improve how teachers and educational developers facilitate such discussions.

Content analysis—the analysis of discussion transcripts based on a pre-defined framework—has become the most used research strategy in relation to online educational discussions (Biasutti, 2017; Wise & Paulus, 2016). Researchers have suggested several frameworks to analyze how students discuss in online educational discussions. Frameworks focusing on argumentation based on Toulmin's argument model are important in the field (Weltzer-Ward, 2011; Wise & Paulus, 2016). Here, I use and discuss a framework by Weinberger and Fischer (2006) that analyzes online educational discussions by focusing on how students construct single arguments (microstructure) and place them in chains of argumentation (macrostructure). This framework builds on Toulmin's (1958/2003) argumentation model.

Several review articles have pointed to the abundance and weak theoretical grounding of frameworks and the need to critically discuss these frameworks (Clark, Sampson, Weinberger, & Erkens, 2007; De Wever, Schellens, Valcke, & Van Keer, 2006; Weltzer-Ward, 2011). Consequently, a critical discussion of how Toulmin's model functions as a means to analyze students' arguments is needed.

The aim of this article is twofold. First, it explores what occurs in online discussions in an introductory philosophy course where competence in argumentation is a specific learning objective, analyzed using the categories of the microstructure and macrostructure of argumentation. Second, it discusses how suitable the categories from Toulmin's model are for such analysis.

Method and description of the empirical context

Analytic framework

What is an argument and how do we analyze argumentation? In its most basic form, an argument is a linguistic expression consisting of at least one part that expresses a claim or statement of opinion. Ideally, another part (or parts) of the argument serves as backing for the claim and eventually a limitation of the claim (Leitao, 2000; Toulmin, 1958/2003). The *processes of argumentation* involve discussants who present claims they find plausible to convey opinions to their audience. If an opponent contests an adduced claim, she/he may be contesting the claim itself and/or the information backing the claim. Sequences of argumentation thus consist of arguments, counterarguments, and replies. Counterarguments may propose new opinions or address the tenability or relevance of the initial claim(s) or their backing.

Weinberger and Fischer (2006) developed a coding scheme to analyze the argumentative construction of knowledge in online educational discussions. Their model comprises four dimensions[1]. The argument dimension—the focus of this article—emphasizes two aspects of argumentation: a) the macro-level of argumentation, based on Leitao (2000), which focuses on sequences of argumentation; and b) the micro-level of argumentation, based on Toulmin (1958/2003), which focuses on the structure of single arguments. In the next section, I present the categories of the microstructure of arguments and the macrostructure of argumentation. In the analysis section, I identify arguments of the different categories.

Analysis of the macro-level of argumentation

The analysis of the macro-level of argumentation focuses on how discussants link posts together as sequences of argumentation. The lack of links between posts shows the discussion lacks a common focus. Table 1 contains an overview of categories of links between arguments and their descriptions.

1. Weinberger and Fischer (2006) include the following in their coding scheme:1) Participation—both overall activity and heterogeneity of interaction. 2) Epistemic dimension—construction of problem space; conceptual space; adequate relations between problem and concept; inadequate relations between problem and concepts; relations between prior knowledge and problem. 3) Argument dimension—a) micro-level, how arguments are backed by evidence (as described above, based on Toulmin's model); and a) macro-level, how a chain of arguments support or criticize a position (arguments, counterarguments, integration of positions/arguments). 4) Social mode of co-construction—articulating thoughts; questioning; coordinating the discussion; integration-oriented consensus building; conflict-oriented consensus building. This results in a complex scheme where coding according to the four dimensions may overlap.

Table 1 Categories of the macrostructure of sequences of argumentation

Category	Description
Argument	Statement put forward to convey an opinion and convince interlocutors
Counterargument	Argument that opposes the preceding argument and proposes an alternative opinion
Integration (reply)	Statement that aims to integrate two arguments
Non-argumentative moves	Questions, coordination, meta-statements

Based on Weinberger and Fischer (2006)

In this article, links between discussion posts are identified. How posts are linked as starting points, counterarguments, or attempts to integrate may indicate the presence or lack of explicit argumentation.

Analysis of the micro-level of argumentation

Toulmin's work on informal logic (Toulmin, 1958/2003) is a standard account of the components and structure of single arguments (Andriessen & Baker, 2014; Inch & Warnick, 2011). According to Toulmin, an argument consists of advancing a *claim*, a proposition the utterer holds true and wants to defend. The claim needs support to be justified and accepted as true. *Data* works as evidence for claims. In many cases, we need something that connects the data to the claim and guarantees the data really gives evidence for the claim. Toulmin calls the premise that guarantees the inference from data to evidence as *backing*. Backing may need some kind of support, called the *warrant*. However, the data, backing, and warrant may support the claim with varying degrees of certainty. The claim may follow with necessity or some degree of probability. Estimating the probability of the claim is part of the line of argument and is called the *qualifier*. The circumstances under which the claim is not true is the *rebuttal*. Figure 1 illustrates the model.

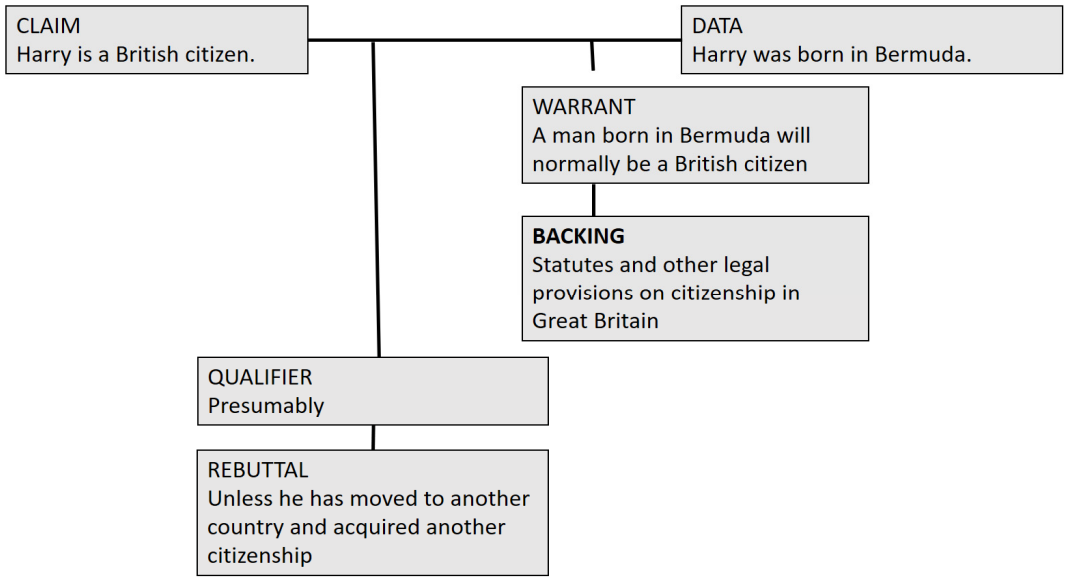


Figure 1 Toulmin's argument model (Based on Toulmin, 1958/2003)

Weinberger and Fischer (2006) based their coding scheme for the microstructure of arguments on a simplified version of Toulmin's model. In the simplified model (Figure 2), data, warrant, and backing are merged under the heading *backing*. The rebuttal and qualifier are merged under the heading *qualification*.

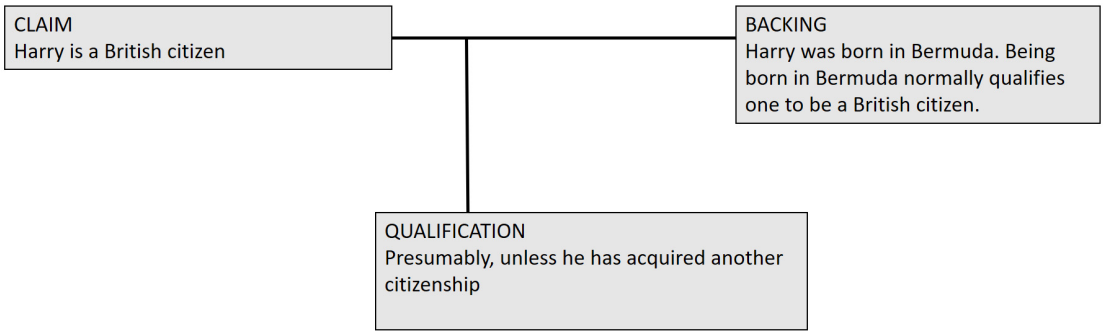


Figure 2 Simplified Toulmin model

The simplified model provides categories to determine whether an utterance contains a claim, whether the claim is supported with backing justifying the claim, and whether the limitations of the claim are explicated. The frequent occurrence of posts and arguments lacking these components indicates weaker argumentation.

Table 2 presents an overview of the categories of the microstructure of arguments.

Table 2 Categories of the microstructure of single arguments

Category	Description	Component
Simple claim	Statement that raises a claim without providing any evidence or limiting the context of where the claim is asserted to be valid and true	CLAIM
Grounded claim	Statement that raises a claim and provides evidence proving the claim but lacks a limitation of the context wherein the claim is valid	CLAIM BACKING
Qualified claim	Statement that raises a claim and limits the context where it is true and valid but lacks evidence proving its truth	CLAIM QUALIFICATION
Grounded and qualified claim	Statement that raises a claim, provides evidence proving the truth, and demarcates the context wherein the claim is valid	CLAIM BACKING QUALIFICATION
Non-argumentative moves	Posts that do not raise a substantial claim about the topic discussed, including questions, suggestions on how to proceed with the discussion, etc.	

Based on Weinberger and Fischer (2006)

Material and empirical context

The analyzed discussions were collected from an online introductory philosophy course. As competence in argumentation was an important learning objective in the course description and the teachers considered the online discussions to be vital in achieving this objective, the material is suitable to investigate students' argumentation.

The primary data source is transcripts from discussions in the LMS. Students were informed about the research on their contributions to discussions, and consent was obtained. At the end of the semester, transcripts of LMS discussions were made and put into a spreadsheet. Posts from non-consenting students were removed, and all names were replaced with aliases. The steps in the analysis of the discussions are described at the end of this section.

In addition to the discussion transcripts, course materials (course readings, video lectures, and instructions) and course descriptions (including a description of the learning

objectives) were collected as secondary data and used to describe the context of the discussions, thus enabling an understanding of the topics discussed. Interviews with the teachers provided the background and further information about the rationale for online discussions as a learning activity and the teachers' evaluation of how such discussions worked as a learning activity. Two researchers conducted a semi-structured focus group interview with the two responsible teachers based on an interview guide disseminated to the interviewees in advance. In the interview, online discussions in similar courses were discussed based on the teachers' experience. The interview was recorded, and some parts were transcribed.

The course was offered as part of the first semester of a BA program in philosophy and as an elective for students from other programs. The course covered an introduction to selected philosophical topics. The teaching and learning activities were video lectures, mandatory course readings, individual written assignments, and mandatory participation in online discussions. All teaching and learning activities occurred via the LMS. Twenty-five students enrolled in the course. The analyzed material is from a discussion in week three of the course. The students were split into two discussion groups of 12 and 13. Each student had to contribute at least two posts and was encouraged to post more.

The topic for the analyzed discussions concerned "questions about the meaning of life." This was chosen to appeal to students in an introductory philosophy course. The central issues were the distinction between objective and subjective criteria for meaningfulness and problems with establishing objective criteria for a meaningful life when the basic foundations for such criteria (such as God) are contested. The instructions for the discussion required the students to:

> [D]iscuss the following assertion: *There are only subjective criteria for a meaningful life.* [...] (Excerpt from instructions in LMS)

Two examples followed the assertion: "Blob" watching brainless TV shows and drinking beer and Sisyphus being forced to push a stone uphill only to see it inevitably roll down (in this example, he was given a pill that made him feel happy). The examples showed lives that could commonly be considered meaningless; however, both characters claimed their lives to be meaningful due to their subjective criteria. The assertion in the instructions closely related to an article (Wolf, 2009) on the reading list and rephrased examples from the article and video lectures. However, the assertion raised a claim that contradicted the article's primary claim. The instructions encouraged students to use concepts and ideas from the article to focus "how you place your arguments" and to find "relevant distinctions between the two examples."

Altogether, 18 students wrote more than 10,000 words (equivalent to 20 standard pages) over a six-day period. This is far more than the mandatory requirement and demonstrates students participated in the discussion enthusiastically. The mode of the discussion was asynchronous. Discussants were not logged on at the same time, and there were time lapses between postings. Some posts are rather long; they address more than one topic and contain several arguments. The discussions stayed focused and did not go off topic. Students employed subject terms, demonstrating a good understanding of the discussed topic. This indicates the students read the course readings and watched the video lectures. Almost all posts were well written on a surface level; the language was clear and readable. Even though the topic may have provoked high-flown thoughts, the discussants maintained a balanced, sober approach in their writing.

In the interview, the teachers confirmed that argumentation and critical thinking were objectives of the course as a whole and particularly of the discussions. The teachers aimed to stimulate: a) activity and dynamics (all participating students engaged in the discussions); b) use of subject terms and knowledge from the course material; and c) reflexivity (providing and evaluating reasons for claims). The teachers claimed that discussions functioned as opportunities to practice using subject concepts and knowledge from the course. The teachers' experience was that in these kinds of discussions in similar courses, students succeeded with rational argumentation only to a limited degree. One interviewee summed it up this way:

> [. . .] one might say that they [the discussion posts] are rich, and that they have a philosophical content. People [students] are sitting there, writing up posts that are more individual reflections than contributions to a discussion, meaning that much of a post is not addressed towards a discussant. (Teacher, introductory course in philosophy)

Nevertheless, even if the teachers found the discussions useful learning activities in terms of practicing subject terms and using knowledge from course readings, they noticed the students succeeded with rational argumentation only to a limited degree. This made it interesting to analyze how students actually constructed their arguments according to the categories of the microstructure of arguments and the macrostructure of argumentation.

Analytic procedure

The analysis of the discussion transcripts comprised several steps. First, macrostructures of argumentation were analyzed. How students referred to the assignment, the initial assertion, and previous posts was identified and visualized. Second, the microstructure of the arguments was identified and categorized. Posts were condensed and paraphrased to make the argumentative structure more transparent. Words and phrases that may indicate a relationship between a claim and backing (conclusion and premises), such as *therefore*, *because*, and *so*, were read with extra caution. Finally, the students' meta-comments about the discussion were carefully read to compare and triangulate their evaluations with the analysis and with the teachers' evaluation stated in the interview.

Analysis

Analysis of the macrostructure of argumentation and students' responses to the assignment

Macrostructures of argumentation describe how discussants link posts together in chains of arguments, counterarguments, and attempts to integrate arguments and counterarguments. According to Leitao (2000), opposition and contrasting opinions are core characteristics of argumentation processes. Figure 3 illustrates how students linked their posts to previous posts. Posts where discussants directly address other discussants by name are categorized as explicit references. Posts where discussants take up specific points without using the name of those they address are categorized as implicit references.

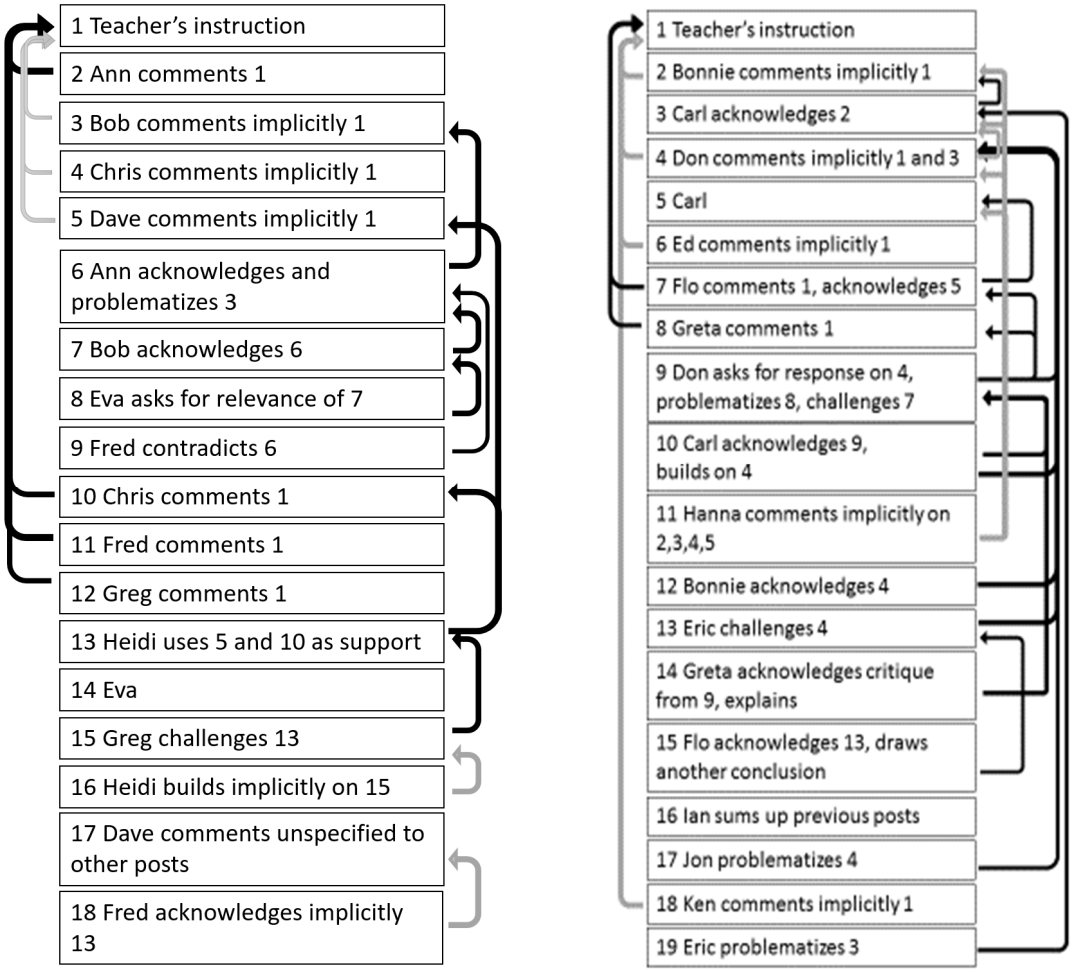


Figure 3 Discussants referring to other posts, groups 1 and 2. Arrows on the left illustrate references to the teacher's instructions; arrows on the right illustrate references between students' posts; black arrows indicate explicit references; and gray arrows indicate implicit references (referring to an argument without mentioning who posted it).

In discussion 1, six posts refer explicitly to posts by other students. Seven posts do not refer explicitly to other posts at all, although they relate to the discussion topic. Four posts refer explicitly to the initial post by the teacher. Discussion 2 shows a somewhat different pattern. Only two posts refer explicitly to the initial post by the teacher. Eleven posts explicitly address posts by other students. Seven posts do not refer explicitly to other posts at all, although they relate to the discussion topic.

How do students place their arguments in relation to previous posts? How do they indicate whether their posts represent counterarguments or attempts to integrate opposing claims? Challenging claims and backing claims are vital for argumentation to achieve improved understanding. According to Littleton & Mercer (2013) and Wegerif (2015), a focus on reasons and justification, challenging and problematizing other views, and asking for clarification, are hallmarks of explorative discussions." Such discussions are suitable to develop new knowledge and better understanding. The examples below show how students referred to each other's posts critically and exploratively.

- "Ann, I do not completely agree with your claim that . . ." (Fred, post 9, group 1)
- "To Heidi, what if we look at the example from a different angle: . ." (Greg, post 15, group 1)
- "Greta, I do not understand your argument [. . .] To Flo who writes that [. . .]. How is this possible? I would claim that [. . .] (Don, post 9, group 2)

About half of the posts have such characteristics. The other half have explicit references to other posts and consist of acknowledging, supporting, or building upon previous posts rather than critiquing or offering contrasting views:

- "Wolf argues, as you, Ann, comment, that . . ." (Bob, post 7, group 1)
- ". . . as Chris and Dave say . . . (Heidi, post 13, group 1)
- "Exciting to read your post, Bonnie." (Carl, post 2, group 2)
- "I think the same as Carl . . ." (Flo, post 7, group 2)
- "Don points to something important here" (Carl, post 10, group 2)

This illustrates the potential for argumentation, as opposition and contrasting (Leitao, 2000) and explorative discussions (Littleton & Mercer, 2013; Wegerif, 2015) were utilized only to a limited degree.

Analysis of students' responses to the instructions for the discussion/discussion task

Thirteen posts refer to the initial post by the teacher stating the discussion topic, the assertion that "[t]here are only subjective criteria for a meaningful life." The assertion relates closely to and contradicts the main message in an article from the course readings.

What does it mean to "discuss an assertion," and what are the criteria to successfully complete such an assignment? "Discuss an assertion" is a commonly used phrase for assignments. Interpreted loosely, it means "elaborate on the topic" or "exchange some ideas" about the topic. A more narrow and analytical interpretation of the task demands that discussants focus on the initial assertion, propose several arguments for and against, evaluate and weigh such arguments, and eventually evaluate the tenability of the initial assertion.

How did the students relate to the initial assertion ("There are only subjective criteria for a meaningful life")? Only a few of the discussion posts related explicitly to the initial assertion. Four posts pointed out that the assertion is contrary to the message in the article (Wolf, 2009) to which the assignment refers. A few posts took an explicit stance regarding the assertion, claiming the assertion is untenable and pointing to its problematic implications. Posts that did not refer explicitly to the initial assertion discussed the more general question about meaning and life. Some of these argued in favor of subjectivity as primordial, implicitly supporting the initial assertion. Other posts discussed the examples and the distinctions between them.

An evaluation of how the discussants discussed the assertion—in the sense of presenting arguments for and against it and of its tenability—indicates the discussion as an assignment failed, as only a few posts addressed the tenability of the initial assertion. However, the discussion was successful as an arena for students to practice using subject terms and to discuss knowledge from course readings by exchanging ideas with peers.

Analysis of the microstructure of arguments

The previous paragraph focused on how discussants related their posts to other posts. Next, we look at how discussants constructed their arguments—that is, the microstructure of arguments. According to established theory (Toulmin, 1958/2003; Weinberger & Fischer, 2006), arguments consist of a *claim* that expresses the utterer's intention or belief he/she wants to convey, the *backing* of the claim that provides the reason to accept the claim, and the *qualification/limitation* of the claim that states the limitations of the claim's validity.

According to Weinberger and Fischer (2006)'s coding scheme, the most sophisticated posts contain an explicit claim, a backing or justification of the claim, and a qualification that limits the claim's validity. Successful argumentation depends on the discussants' ability to put forward such arguments. Therefore, arguments that lack *backing*, *qualification*, or both represent weaker modes of argumentation.

After condensing the material, all the argument types described in Weinberger and Fischer (2006)'s coding scheme were identified. Examples of several types of arguments are shown below.

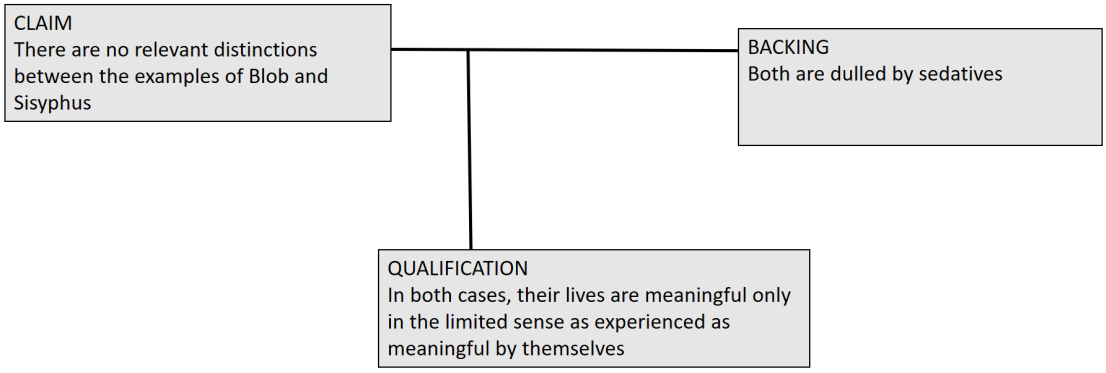


Figure 4 Claim with backing and limitation

Figure 4 illustrates a complete argument containing a claim, backing, and rebuttal. The utterer responds to a question raised in the teacher's instruction about whether the two examples, Blob and Sisyphus, are the same in terms of meaningful lives or whether there are relevant distinctions. The utterer's main claim is that there are no relevant distinctions. This is backed by pointing to what the two examples have in common. Further, the utterer draws attention to the limitation of what can be deduced from the examples. Even if the examples of Blob and Sisyphus have no relevant distinctions between the meaningfulness of their lives, their lives' meaningfulness is limited only by the degree to which they experience them as meaningful. However, this last component of the argument may also be interpreted as a separate claim responding to the more general discussion about subjective versus objective criteria for meaningful lives. According to this interpretation, other parts of the post may serve as backing and/or qualification.

However, this kind of argument containing a *claim*, *backing*, and *qualification* is atypical for the material. This category of argument represents fully developed, complete arguments and thus the ability to practice argumentation, are rare in the material. More incomplete variants and less-developed arguments are more frequent. Examples of such arguments contain a claim but lack qualification and/or backing. Below are examples of these categories.

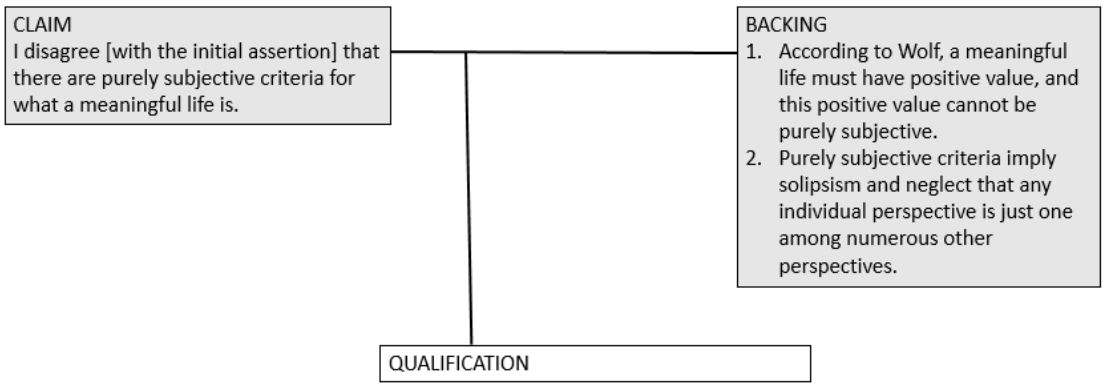


Figure 5 Claim with backing but without qualification

The utterer disagrees with the initial assertion. As backing for this claim, she/he presents arguments based on the course readings. The first backing is that Wolf (2009) states positive value is a prerequisite for a meaningful life and that such value cannot be purely subjective. The second is that purely subjective criteria imply solipsism, thus neglecting other perspectives. No limitation of the claim or the specific circumstances under which the claim or backing is valid are specified. For the first part of the backing, a limitation may be that one needs to accept Wolf's argument as authoritative if this backing is to have any weight. If a discussant does not accept Wolf's line of reasoning, she/he will probably be unconvinced by the first utterance in the argument's backing. The lack of an explicated qualification in the argument may indicate the utterer does not see the limitation of the backing or recognize how the argument may be refuted.

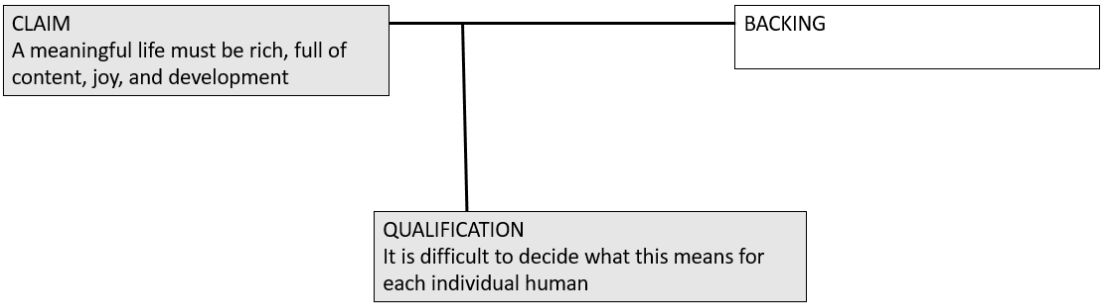


Figure 6 Claim with qualification but without backing

Figure 6 illustrates one among several variants of incomplete arguments. The utterer claims that a meaningful life must contain some specific elements. The post contains no attempts to back or justify this claim. Perhaps the utterer takes the claim to be self-evident or intends to build on something that was previously backed in the discussion. Because the utterer did not explicate how she/he intends to back the claim, we as the audience (fellow discussants or evaluators of the discussion) have no opportunity to considerer any backing, but observe that backing is omitted. Nevertheless, in this post, the utterer explicates a limitation to the claim. Again, it is not made explicit by the utterer whether this represents a qualification and thus a limitation of the claim's extension, or a problematization of the claim and thereby a counterargument to the claim.

A category of arguments that Weinberger and Fischer (2006) described as less advanced and that consequently contribute the least in a discussion are arguments that raise a claim

without providing justification or backing or any qualification that limits the validity of the claim. Such arguments are quite common in the material. Figure 7 illustrates a claim that lacks both backing and qualification.

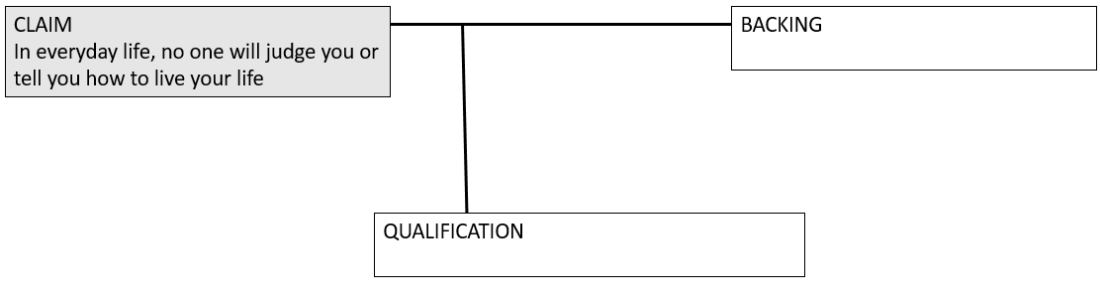


Figure 7 Claim without backing or qualification

One example is shown above. One discussant states that (normally) people do not judge or tell others how to live their lives. In the post, the utterer continues by uttering that *no one will tell you whether or not you have a meaningful life*. The context of this statement is a sequence in the discussion where another discussant claimed that judging each other's life implicitly indicates that some kind of (objective or at least common) criteria for meaningful lives are available. This statement functions as a counterargument. However, concerning the microstructure of arguments, this claim stands out as an unsupported and unjustified claim.

In summary, condensing students' posts and categorizing them according to Weinberger and Fischer (2006)'s microstructures of arguments showed that all types of arguments are found. Most frequent are arguments that raise a claim without providing backing or limitation. Complete arguments (containing an explicit claim, backing, and qualification) are underrepresented. This coincides with previous research (Noroozi, Weinberger, Biemans, Mulder, & Chizari, 2013; Tsai & Tsai, 2014; Wise & Hsiao, 2018).

Students' meta-comments about argumentation

Posts where students explicitly commented on the discussion or their own understanding were read with extra caution. Typically, such passages contained apologies for their problems understanding the topic and described difficulties with clearly articulating their ideas. In some passages, students also commented on how the discussion took form:

> Just a little objection: Can we be better at discussing WITH each other and using one another's points, either criticize/slaughter them or back them up? With some exceptions, this discussion is mostly our own posts with our/some opinions about the topic; however, without addressing the disagreements between us, we can't argue our points well.

Interestingly, the student cited here recognized that the discussants' comments on each other's posts remained on a surface level, even when they addressed each other explicitly. However, without addressing each other's claims or arguments, no real discussion occurred. Another student responded, explaining that he found it demanding to take a stance on the questions discussed:

> We want . . . first and foremost to convince ourselves about what we mean about these things. Therefore, maybe for some of us, it is enough to say something about where you place yourself in relation to the text, rather than having strong opinions against (or strong support for) what the other discussants say.

A third student entered the conversation. His comment may be interpreted as stating that he found the topic challenging and that discussions with peers are stimulating and inspire his own thinking:

> It is easy (at least for me) to get entangled in such thoughts, and so I understand why philosophy is best done in the company of others ;-).

This passage is a view into the students' difficulties understanding the discussed subject and their quest to master discussing these topics. The students' reflections show both understanding of challenges according argumentation and challenges according understanding of subject topics. Yet, analyzed by using microstructures of arguments, such posts will be classified as "non-argumentative moves" (posts that do not raise a substantial claim about the topic discussed, including questions, suggestions on how to proceed the discussion, etc.; see Table 2). This may indicate a deficiency with Weinberger and Fischer (2006)'s coding scheme focusing on the micro- and macrostructure of arguments—namely that students' self-reflection is disregarded. These reflections are interesting, valuable parts of the material. Yet, it is interesting to note that students identify similar shortcomings in the discussions, as identified by analyzing according to the macrostructure, namely that the discussants did not address each other. The coinciding between students' meta-comments on this theme and the analysis of the discussion transcripts represents a form of informant validation of the analysis, which is also supported by the teacher interview.

Summing up and implications

According to the first research question of how students construct and place arguments in an online educational discussion, analyzed by the categories of the microstructure of arguments and the macrostructure of argumentation, the analysis of the macrostructure of argumentation showed that even though the students showed good knowledge of course readings, they commonly wrote posts without clear reference either to other students' posts or to the initial assignment. Instead of constructing lines of argumentation consisting of arguments, counterarguments, and the integration of arguments, students posted their thoughts without any clear and explicit link to other discussants' arguments. Thus, the potential for expanding understanding by contrasting points of view is fulfilled only to a limited degree. Further, responding to the assignment "discuss the following assertion . . .", the students seem to interpret "discuss" as an invitation to write whatever they thought about the topic. More elaborated responses to the assignment are rare, such as the presentation of arguments for and against the initial assertion, or taking a stand regarding the initial assertion based on some kind of evidence or argument, or critiquing other discussants' stands regarding the initial assertion by scrutinizing the backing they offer. Further, the arguments students constructed tended to be incomplete according to Toulmin's model, lacking backing and/or limitation of the claim's validity.

The potential for generalization based on a small amount of material is limited. Yet, previous studies had similar findings—students' argumentative level in online educational discussions is commonly weak. Interestingly, the students' meta-comments and the teachers' general impression of online discussions coincide with the analysis in this article; even if the argumentative level is weak, the discussions are both interesting and in some ways are worthwhile learning activities. Nevertheless, the analysis performed with the categories of the arguments' micro- and macrostructure provide a more detailed, profound picture of what occurs in such discussions than was found in the teachers' evaluation or in the students' meta-comments.

This leads to the second research question: How suitable are the categories used in the analysis? As shown above, a more profound picture of what occurs in the discussions is possible based on the categories. The categories are powerful lenses with which to identify the strengths and weaknesses in argumentation. Yet, analyzing the microstructure of arguments requires careful interpretation and involves reconstructing more or less disguised argumentative structures in students' writings to see how these structures fit into the established classification. It might be tempting to over-interpret and take a far too charitable approach, which might result in adding too much sophistication to the argumentation. Another risk is employing too little sensitivity, thereby missing some of the argumentative qualities. This difficulty is identified in previous research (Erduran, Simon, & Osborne, 2004, p. 919) and relates to clarifying which parts of a linguistic expression count as a claim, backing, or qualification. In natural language, like the students' posts, it is rather implicit which parts (are meant to) serve as a claim, backing, or qualification, and the components of arguments (claim, backing, and qualification) are presented in contingent order. Further, a statement in the discussion may function as a claim or as a backing for a recently posted claim. The discussants wrote without any attempt to fulfill the argument pattern prescribed by Toulmin. Therefore, categorizing their writings according to this model was not straightforward. In some cases, the relationships between posts or between argument parts might have been rather implicit. These demanded careful interpretation when analyzing and categorizing the argument structures. Therefore, even if this argumentation framework represents a sound approach to analyzing argumentation, careful consideration is necessary when using the framework to analyze natural language. Moreover, it is worth noting that students' self-reflection, as presented above and which may be important for students' development of argumentative and substantial competence, is not acknowledged by Weinberger and Fischer's categories.

In summary, students participated in the online discussion enthusiastically, yet the quality of their argumentation is weak according to this model. Micro- and macro-levels of argumentation are useful categories to analyze what occurs in discussions. Analysis based on the categories provides a richer, more detailed picture of students' argumentation than teachers' and students' evaluations. Nevertheless, analysis using these categories is not straightforward and demands careful interpretation.

What can be learned from this analysis of the transcripts of online discussions among students in an introductory philosophy course? It is striking that the students demonstrate a weak understanding of what "argumentative coherence" means. In the discussion, they fail to back their claims and to address their arguments to discussants' claims. This is probably because they need to elaborate an understanding of the qualities of argumentation related to the micro- and macrostructures of argumentation. The students seem to be busy trying to grasp the concepts used in the course literature to discuss "subjective and objective criteria for a meaningful life". Of course, this is worthwhile; however, it seems too daunting a challenge to both manage complex subject knowledge and subject terms and simultaneously explore how to argue coherently.

For teachers who set up online discussions as learning activities to promote deeper learning and/or competence in rational argumentation, the micro- and macrostructure of argumentation may be elusive categories. To request that students explicate what their claim is, how it is backed and limited, and which of the discussants' arguments they address, may enhance the quality of their postings. Without explicitly understanding basic features of rational argumentation, students are left to discover the craft of argumentation by themselves. Learning is therefore a result of osmosis—knowledge and skills transmitted by being

in a salient milieu or due to serendipity. Learners grasp the craft and categories of argumentation by stumbling upon them or inventing them themselves. Learning something inductively and implicitly by osmosis or serendipity may work; however, by learning in such ways, students miss understanding explicit and precise concepts about argumentation and consequently lack precise instruments to evaluate and discuss argumentation with their peers.

In the introduction, I sketched the idea that students' comprehension of a subject is reflected in the quality of the arguments they are able to present about that subject. Mastering the craft of argumentation is thus a powerful tool to learn a subject. In mastering the ability to identify claims, the backing and limitation of claims are key competencies that enable learners to scrutinize and evaluate the tenability of knowledge claims.

References

Andriessen, J., & Baker, M. (2014). Arguing to learn. *The Cambridge Handbook of the Learning Sciences*, pp. 439–460 Cambridge University Press.

Arum, R., & Roksa, J. (2011). *Academically adrift: Limited learning on college campuses*: University of Chicago Press.

Biggs, J., & Tang, C. S.-k. (2007). *Teaching for quality learning at university: What the student does*. Maidenhead: McGraw-Hill/Society for Research into Higher Education & Open University Press.

Cahill, A. J., & Bloch-Schulman, S. (2012). Argumentation Step-By-Step. *Teaching Philosophy, 35*(1), 41–62. https://doi.org/10.5840/teachphil20123514

Clark, D., Sampson, V., Weinberger, A., & Erkens, G. (2007). Analytic Frameworks for Assessing Dialogic Argumentation in Online Learning Environments. *Educational Psychology Review, 19*(3), 343–374. https://doi.org/10.1007/s10648-007-9050-7

De Wever, B., Schellens, T., Valcke, M., & Van Keer, H. (2006). Content analysis schemes to analyze transcripts of online asynchronous discussion groups: A review. *Computers & Education, 46*(1), 6–28. https://doi.org/10.1016/j.compedu.2005.04.005

Dewey, J. (1998). *How we think : a restatement of the relation of reflective thinking to the educative process*. Boston: Houghton Mifflin.

Ennis, R. H. (1989). Critical thinking and subject specificity: Clarification and needed research. *Educational researcher, 18*(3), 4–10.

Erduran, S., Simon, S., & Osborne, J. (2004). TAPping into argumentation: Developments in the application of Toulmin's Argument Pattern for studying science discourse. *Science Education, 88*(6), 915–933. https://doi.org/10.1016/j.compedu.2005.04.005

Facione, P. A., & Facione, N. C. (2007). Talking Critical Thinking. *Change: The Magazine of Higher Learning, 39*(2), 38–45.

Inch, E. S., & Warnick, B. (2011). *Critical Thinking and Communication: The Use of Reason in Argument*. Boston: Allyn & Bacon.

Kovanović, V., Gašević, D., Joksimović, S., Hatala, M., & Adesope, O. (2015). Analytics of communities of inquiry: Effects of learning technology use on cognitive presence in asynchronous online discussions. *The Internet and Higher Education, 27*, 74–89.

Leitao, S. (2000). The potential of argument in knowledge building. *Human Development, 43*(6), 332–360. https://doi.org/10.1159/000022695

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

Loncar, M., Barrett, N. E., & Liu, G.-Z. (2014). Towards the refinement of forum and asynchronous online discussion in educational contexts worldwide: Trends and investigative approaches within a dominant research paradigm. *Computers & Education, 73*(0), 93–110. Retrieved from https://doi.org/10.1016/j.compedu.2013.12.007

Marton, F., & Säljö, R. (1976). On qualitative differences in learning: I – Outcome and process. *British Journal of Educational Psychology, 46*(1), 4–11. Retrieved from https://doi.org/10.1016/j.compedu.2013.12.007

Noroozi, O., Weinberger, A., Biemans, H. J. A., Mulder, M., & Chizari, M. (2013). Facilitating argumentative knowledge construction through a transactive discussion script in CSCL. *Computers & Education, 61*(0), 59–76. doi:http://dx.doi.org/10.1016/j.compedu.2012.08.013

Rourke, L., & Kanuka, H. (2009). Learning in Communities of Inquiry: A Review of the Literature. *Journal of Distance Education, 23*(1), 19–48. http://www.jofde.ca/index.php/jde/article/view/474.

Salmon, G. (2013). *E-tivities: the key to active online learning*. New York: Routledge.

Schindler, L. A., & Burkholder, G. J. (2014). Instructional Design and Facilitation Approaches That Promote Critical Thinking in Asynchronous Online Discussions: A Review of the Literature. *Higher Learning Research Communications, 4*(4), 11–29. https://doi.org/10.18870/hlrc.v4i4.222

Siegel, H. (1988/2013). *Educating reason*: London: Routledge.

Thomas, J. (2013). Exploring the use of asynchronous online discussion in health care education: A literature review. *Computers & Education, 69*(0), 199–215. Retrieved from https://doi.org/10.1016/j.compedu.2013.07.005

Toulmin, S. (1958/2003). *The uses of argument*. Cambridge: Cambridge University Press.

Tsai, P.-S., & Tsai, C.-C. (2014). College students' skills of online argumentation: The role of scaffolding and their conceptions. *The Internet and Higher Education, 21*, 1–8. Retrieved from https://doi.org/10.1016/j.iheduc.2013.10.005

Wegerif, R. (2015). Toward Dialogic Literacy Education for the Internet Age. *Literacy Research: Theory, Method, and Practice, 64*(1), 56–72.

Weinberger, A., & Fischer, F. (2006). A framework to analyze argumentative knowledge construction in computer-supported collaborative learning. *Computers & Education, 46*(1), 71–95. Retrieved from http://www.sciencedirect.com/science/article/pii/S0360131505000564. doi:http://dx.doi.org/10.1016/j.compedu.2005.04.003

Weltzer-Ward, L. (2011). Content analysis coding schemes for online asynchronous discussion. *Campus-Wide Information Systems, 28*(1), 56–74.

Wise, A. F., & Hsiao, Y.-T. (2019). Self-regulation in online discussions: Aligning data streams to investigate relationships between speaking, listening, and task conditions. *Computers in Human Behavior, 96*, 273–284. https://doi.org/10.1016/j.chb.2018.01.034

Wise, A. F., & Paulus, T. M. (2016). Analyzing learning in online discussions. In Andrews, R & C Haythornthwaite (eds.): *The SAGE Handbook of E-learning Research*. SAGE Publications Ltd

Wolf, S. (2009). The meaning of lives. In Perry, Bratman, & Fisher (Eds.), *Introduction to Philosophy*. New York: Oxford University Press.

Zheng, B., & Warschauer, M. (2015). Participation, interaction, and academic achievement in an online discussion environment. *Computers & Education, 84*, 78–89. https://doi.org/10.1016/j.compedu.2015.01.008

Toulmin's Argument Model Used to Analyze Critical Thinking in Online Educational Discussions – An Overview and Critical Evaluation

Jens Breivik
Associate Professor
Jens.breivik@uit.no
Phone: +47 951 15 241
Department of Education
UiT—The Arctic University of Norway

N-9037 TROMSØ

Norway

Keywords:

Online educational discussions; argumentation; content analysis; Toulmin's argument model

Abstract

Educators and researchers strive to explain how participation in online educational discussions influences students' argumentation and, thus, their learning. To do this, researchers need a framework with which to analyze the quality of argumentation. Scholars reviewing the field have identified more than 50 frameworks for the content analysis of online educational discussions. However, a lack of consensus about which frameworks best assess the quality of argumentation in online educational discussions hampers opportunities to build on previous research and compare findings. Frameworks and coding schemes based on Toulmin's argument model, which focuses on the micro-structure of arguments, are held to be among the most prominent in this research field. These frameworks suggest that researchers can assess rational argumentation as the ability to produce complete arguments, which contain a claim, the backing of this claim, as well as possible limitations and rebuttals regarding this claim. First, this article presents an overview of how frameworks based Toulmin's model are used in the research field to analyze online educational discussions. The perception that such frameworks are used extensively will be discussed in the article. Second, the article evaluates the adequacy of such frameworks by using Toulmin categories to analyze excerpts of arguments from a discourse about the Earth being flat or spherical. The article fills a research gap identified in previous review studies, i.e., it considers and problematizes the adequacy of important research instruments in the field. The article concludes with the assertion that counting the presence of the components of Toulmin's argument seems to be an inadequate approach to analyzing the argumentative quality of students' contributions in online educational discussions.

Keywords

Online educational discussions; argumentation; content analysis; Toulmin's argument model

Research Questions

1. How is Toulmin's argument model used in research on online educational discussions?

2. How adequate is this model for the analysis of argumentation and learning in online educational discussions?

How the Research in this Paper Advances the Field

The article contributes to methodological debates about how argumentation in online educational discussions can be analyzed. The analytic strategy of counting the presence of argument components based on Toulmin's model is discussed and problematized.

1. Introduction

Argumentation is a key competence in almost every academic subject, and a growing body of literature points to links between the quality of argumentation and that of learning (Andriessen & Baker, 2014; Nussbaum, 2011). Online educational discussion is a learning activity commonly used to enhance students' capacity to present and evaluate rational arguments. Online discussions include affordances to mediate student interactions in a flexible mode, thus allowing students to write their own arguments and consider those of their peers at their own pace.

Both educators and researchers seek to understand how discussion and collaboration influence students' ability to think critically, to present and evaluate rational arguments, and, ultimately, how these factors affect their learning. Questions in the research field typically deal with the best methods for facilitating online discussions among students and the outcomes of such discussions. Researchers employ several research methods to analyze students' argumentation in online educational discussions, with content analysis being the most frequently deployed (Gao et al., 2013; Wise & Paulus, 2016). Content analysis consists of analyzing discussion transcripts based on a predefined framework and coding scheme, with the aim of identifying and analyzing crucial aspects of students' writings. This analysis serves as both an observational study that provides a window into cognitive and social processes and an analysis of the products of such processes.

Previous research has suggested several theoretical frameworks for the content analysis of online educational discussions. In a comprehensive review, Weltzer-Ward (2011) identified more than 50 frameworks and coding schemes used in research on learning in online discussions. Researchers continue to develop and propose new frameworks to analyze student learning through online discussions (see, e.g., Biasutti, 2017), with much of the focus being on the students' critical thinking and rational argumentation.

Several scholars in the field have pointed to the need to critically discuss both the validity of the frameworks and the links to their theoretical backgrounds (Clark et al., 2007; De Wever et al., 2006; Noroozi et al., 2012; Nussbaum, 2011). More than a decade ago, De Wever et al. (2006, p. 6) raised a concern that remains pertinent in this research field:

> There are questions about the coherence between theoretical base and the operational translations of the theory in the instruments. Instruments are hardly

> compared or contrasted with one another… [and] validity of the instruments [is] limited.

De Wever et al. observed that, despite the growing research literature on the content analysis of online educational discussions and methodological approaches, there was a persistent need for thorough discussions about the suitability of proposed coding schemes. This observation was echoed by Clark et al. (2007) in their review of several frameworks used for the content analysis of argumentation in online educational discussions:

> This variety of perspectives leads to the obvious conclusion that it is insufficient for researchers to say "we measured the quality of argumentation" or "we successfully supported argumentation." Researchers need to specify the theoretical interpretation of argumentation underlying their analytic methods in order to facilitate communication and comparison in relation to other research in the field. Clearly there are multiple theoretical perspectives and aspects of argumentation worth fostering. Clarity of communication regarding theoretical commitments and pedagogical goals is therefore critical in terms of environment design and analytic frameworks. (Clark et al., 2007, pp. 367–370)

According to Clark et al., it appeared that argumentation was considered a phenomenon, even without the need for further clarification. They, therefore, emphasized that argumentation is not a single, easily conceptualized subject. Similar to De Wever et al. (2006), they called for the need for more thorough discussions concerning analytic frameworks and their theoretical backing.

Validity concerns the quality of inferences that researchers draw from data. A lack of research instruments (i.e., coding schemes) to adequately bridge data and theoretical constructs makes it difficult for researchers to draw valid inferences and move beyond the process of data gathering (De Wever et al., 2006). A lack of consensus as to which frameworks best assess argumentation hampers the efforts to build on previous research. Therefore, it is urgent to discuss how adequate these frameworks and coding schemes are for conceptualizing students' learning of argumentation skills when applied to the content analysis of online educational discussions.

Despite concerns about the theoretical bases and frameworks for data interpretation, the research field is growing. In a handbook chapter, Wise and Paulus (2016) pointed to a commonly used group of analytic frameworks and coding schemes:

> One of the most extensively utilized group of models for conceptualizing learning in online discussions is grounded in argumentation as a desired form of academic talk. Models for thinking about argumentation in online discussions are often based on Toulmin's model which conceptualized *individuals' argumentative moves* as claims (possibly warranted with grounds), rebuttals and qualifiers. (Wise & Paulus, 2016, pp. 273-274)

Similarly, Weltzer-Ward (2011, p. 69) found that "variations on Toulmin's (1958) argument framework which provides a theoretical basis for describing argument construction have also been extensively employed [as the basis for content analysis coding schemes for online discussions]." Other reviewers have reached a similar conclusion (Noroozi et al., 2012; Nussbaum, 2011). However, neither Wise and Paulus (2016) nor Weltzer-Ward (2011) provided a thorough review or discussion of Toulmin-based frameworks. In a review of the research on argumentation in STEM education, Erduran et al. (2015, p.11) noted "that there may be qualitative variations in the way that researchers have adapted his [Toulmin's] framework." However, their review did not delve into these variations. Furthermore, Noroozi et al. (2012) commented on the influence of Toulmin's model and discussed the related constraints and affordances as a means to analyze discussions. Among their conclusions was that future reviews of the field should analyze how researchers focus on and operationalize argumentation (p. 101). Other scholars have also criticized the use of Toulmin's model for the analysis of argumentative quality (Andriessen & Baker, 2014; Nussbaum, 2011). Consequently, a discussion about how Toulmin-based frameworks contribute to our understanding of critical thinking, rational argumentation, and learning in online educational discussions is urgent.

In addition to using Toulmin's model, researchers have employed several strategies for analyzing the quality of argumentation. These analytic strategies typically focus on different aspects of argument quality and include the discussants' change of opinion (Nussbaum et al., 2007); conceptual quality, which evaluates the acceptability of an argument's evidence (Clark & Sampson, 2008); students' use of theory in argumentation (De Wever et al., 2007); the macro-structure of argumentations, i.e., their interaction patterns and argumentation sequences (Leitao, 2000; Weinberger & Fischer, 2006); and the epistemic nature of students' reasoning (Duschl, 2007).

This article presents an overview of how frameworks employing the Toulmin model are used in the research field. Furthermore, the article discusses whether these frameworks and coding schemes enable valid inferences to be made about critical thinking and rational argumentation in online educational discussions. Two research questions guide the article:

1. How is Toulmin's argument model used in research on online educational discussions?
2. How adequate is Toulmin's argument model for the analysis of argumentation and learning in online educational discussions?

By tackling these questions, the article contributes to theoretical and methodological debates on how argumentation can be analyzed and how student activity in online discussions can be assessed.

1.1. Toulmin's Argument Model

Toulmin (2003/1958) proposed a model for analyzing the microstructure of arguments as a way of challenging the traditional conception of logic as a formal and a priori discipline. Aristotle's thoughts on rational thinking have guided Western scholars since antiquity, a tradition that states that rational thinking consists of drawing valid conclusions from premises. The paradigmatic form of an argument is syllogism, which consists of a major premise, a minor premise, and a necessary conclusion drawn from the premises:

Major premise:	All humans are mortal.
Minor premise:	Socrates is a human.
Conclusion:	Socrates is mortal.

According to this view, the validity of an inference—from premises to conclusion—is a function of the argument's form. Every deduction—from premises to conclusion—based on a logically valid form will always be valid.

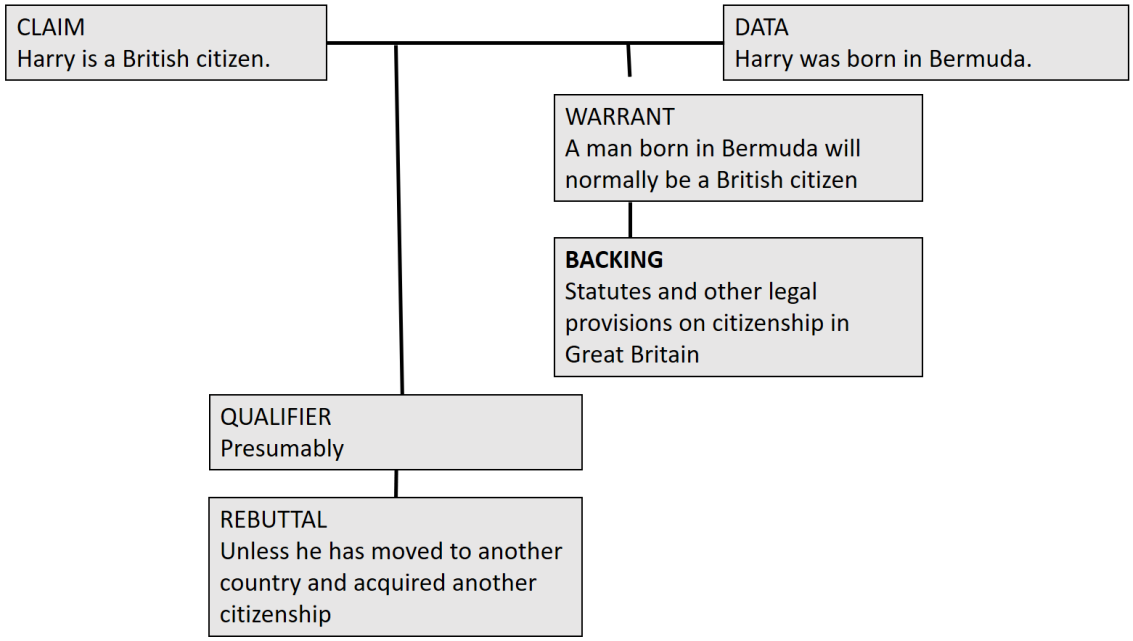
This conception of rationality and reason as applications of formal logic has been challenged, leading to a re-evaluation of how we should educate students to develop competence in rational argumentation. According to Toulmin (2003/1958), real-life argumentation often involves meeting premises and lines of reasoning that may lead us to conclusions that are not necessarily, though are probably, true or false. Probability may be a function of premises that are considered more or less true or an inference from premises to a

conclusion that is not deductively necessary but, rather, inductive and contingent. Toulmin's work is part of a contemporary orientation toward this informal logic.

INSERT FIG. 1.

Figure 1

Toulmin's Argument Model



In Toulmin's model (Figure 1), an argument consists of putting forward a *claim*, which needs support in order to be justified and accepted as true. *Data* work as evidence for claims. In many cases, we would need information that bridges data to the claim and guarantees that the data provide evidence for the claim. Toulmin names this form of premise, which guarantees the inference from data to claim, *backing*. Backing may need some type of support, which is called a *warrant* in this model. However, the data, warrant, and backing may support the claim with varying degrees of certainty. The claim may follow with necessity or a degree of probability. The probability estimate of the claim is also part of the argument and is called a *qualifier*. For some arguments, there is also information indicating the circumstances under which the claim is not true. Such information is called the *rebuttal*. Figure 1 illustrates the relations between the different components of an argument.

2. Methods

This section presents the methodological procedures for gaining an overview of how Toulmin's model is used in the research field and for critically evaluating how adequate coding schemes based on the model are used for analyzing the general quality of argumentation in online educational discussions.

2.1. Methodological approach to research question 1

The starting point in addressing the first research question is the assumption that the Toulmin model is widely used as an analytical tool in researching online educational discussions (Weltzer-Ward, 2011; Wise & Paulus, 2016). How is the model used, and what do the research corpora look like?

For this purpose, literature searches were performed through the EBSCOhost search engine in the Education Research Complete (ERC) and Education Research Information Center (ERIC) databases. These databases were selected due to their wide coverage of literature on educational research. The searches were limited to studies written in English and published in peer-reviewed academic journals. Theses, book chapters, and conference papers were excluded. The first literature search included the keywords "online, educational discussions" and "Toulmin" (see Table 1 in the Appendix for the complete search strings and procedures). Studies were scanned by the author to identify papers that employed Toulmin's model for the analysis of transcripts from online educational discussions. Studies focusing on analyses of classroom dialogue and other non-digital settings were excluded. This criterion excluded influential texts that have used Toulmin's model to analyze off-line settings. Studies that did not analyze discussion transcripts or promote Toulmin's model for such analyses were also excluded.

The search retrieved a very small number of published studies (11), of which only one (Clark & Sampson, 2007) fulfilled the inclusion criteria. The remainder of the articles found in this search focused on discussions in non-digital settings (9) or employed other methods of data analysis, such as surveys and a Delphi study (1). This is far fewer than what one would expect in relation to a framework that is used extensively in the research field. Even well-known, frequently cited studies, such as those by Weinberger and Fischer (2006) and Clark and Sampson (2008)—which Wise and Paulus (2016) reference as examples of important

studies of this kind—were not found among the search results. One possible explanation for the low number of publications produced by the search lies in the standard procedure for a database search, which searches publication titles, abstracts, and keywords. This does not necessarily retrieve all relevant publications. Research studies may well have employed the Toulmin model as an analytical tool without mentioning the phrase "Toulmin" in their titles, abstracts, or keywords.

Another strategy to determine whether the Toulmin's model has been used in the research field is to take as a starting point studies identified by a previous review study. In a comprehensive review of research on ABCSCL (Argument Based Computer Supported Collaborative Learning), Noroozi et al. (2012) identified 89 studies published between 1995 and 2011, which analyzed student argumentation in online educational discussions, some of which employed Toulmin's model. For the present study, the 89 studies identified by Noroozi et al. (2012) were categorized according to their use of Toulmin's model. The result showed that out of 89 the studies, 20 employed Toulmin's model to analyze student argumentation. These studies are included in Table 2 in the Appendix.

To balance the number of pre-2011 studies identified via the review by Noroozi et al., the list of included articles (Appendix, Table 2) was supplemented with a broad search using the keywords "online, educational discussions" and "argumentation" (truncated, see Table 1 for a complete search string). The search, limited to 2018-2020, retrieved 61 studies. "Argumentation" (truncated) in the search string retrieved studies that focused substantially on argumentation as an aspect of learning as well as studies in which the word "argument" (or similar) was used to describe the arguments and claims in the text. Of the 61 studies identified in this search, a relevance scan retrived six studies that analyzed student argumentation in online discussions based on Toulmin's model. These are included in the Appendix, Table 2. The search was limited to post-2018 in order to make both the search and analysis manageable. However, this entailed a lacuna for studies published between 2011 and 2018 (except those identified by hand searches; see the next paragraph). Nevertheless, this number of studies served as a basis to identify variation in the use of Toulmin's model. The analysis of the post-2018 studies showed the same pattern as the remainder of the included studies.

Finally, the list of included articles (Appendix, Table 2) was supplemented by a hand search, which involved tracing literature references and publications in the researcher's personal literature database. The use of several search strategies triangulated the literature

search. The search was closed when additional studies from the second search, searching backwards until 2018, saturated the patterns identified in the previous retrieved studies. Table 1 presents the procedures for several literature searches, including the search strings and results. The included studies are presented in the Appendix, Table 2.

INSERT TABLE 1

Table 1

Literature Search: Procedures and Results

	Keywords	**Search phrase**	**Retrieved publications** (October 2019)	**Publications included after relevance scan**
Search 1	Online, educational discussion AND Toulmin	Toulmin*	11	1 This study is also included in those identified by Noroozi et al. (2012)
Search 2, 2018 and later	Online educational discussions AND argumentation	Argu*	61	5
Common search phrases for both searches		"Online discussion*" OR "Computer mediated communication" OR "CMC" OR "Asynchronous discussion*" OR "Asynchronous communication" OR "discussion group*" AND Education* OR Learning OR Teaching		
Articles reviewed by Noroozi et al. (2012)			89	20
Hand search (Articles from researcher's personal archive, retrieved by following references in other journals and extensive reading of the research literature. Articles found by other searches are not listed in this category.)				15

2.2. Methodological approach to research question 2

The second research question scrutinizes the adequacy of Toulmin's model as a means for the analysis of online educational discussions. Toulmin's argument model has been central to studies of argumentation and the development of informal logic. When researchers use the model to analyze online educational discussions, they commonly justify their choice by pointing to the recognition Toulmin has received. However, there is a need to answer the critical question of whether the model is adequate for the analysis of argumentation. To answer this question, one needs appropriate methods to evaluate the model's adequacy for analyzing rational argumentation.

One step is to condensate the key idea in this approach to measuring argument quality. To anticipate, when researcher use the Toulmin model for the analysis of argumentation, what they actually do is to count the presence of argument components, assuming that a higher number of components represents higher argument quality. Explicitly formulating this key idea contributes to the transparency of the approach, thus opening it up to scrutiny.

Thought experiments are one way of exploring the possible consequences of a principle, a set of categories, etc., by imagining how they will behave in a real-world application (Maxwell, 2013, pp. 68–72). This approach was conducted in the current study by testing examples of how the presence of argument components functions as an indicator for argument quality. In Section 4 of this article, examples from a debate concerning the flat Earth theory were chosen to test how the Toulmin model served as a measurement for the analysis of argument quality. One aim is to check whether such analyses enable false positive or false negative measurements.

3. Analysis: How is Toulmin's Argument Model Used in Researching Online Educational Discussions?

The first literature searches (see Table 2) revealed that Toulmin's influence on the research field is less clear than Weltzer-Ward (2011) and Wise and Paulus (2016) assumed and that the claim that "[m]odels for thinking about argumentation in online discussions are often based on Toulmin's model" (Wise & Paulus, 2016, p. 273) should be scrutinized. The subsequent searches provided a richer picture. The findings from all the searches were compiled in a list of 40 studies and are presented in Appendix, Table 2.

A common research strategy for the analysis of online educational discussions in the included studies is to consider argument quality, analyzed by the Toulmin model, as a dependent variable, influenced by factors such as facilitation and scaffolding, student perceptions, etc. This includes measuring the level of argument quality and determining the statistical relations between independent variable(s) and dependent variable(s) of argumentative quality. Independent variables include different scaffolding strategies, students' gender, self-regulation, reading behavior, etc.

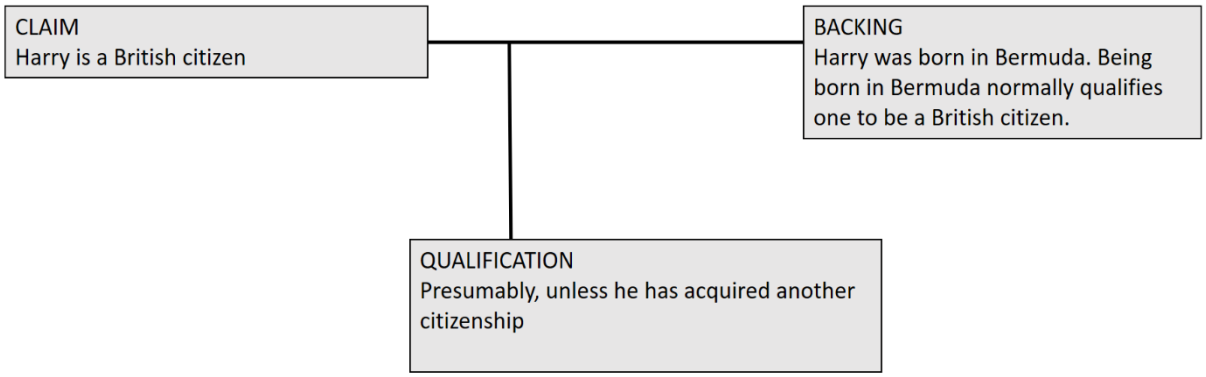
When using Toulmin's model, researchers code the occurrence of argument components (*claim, data, warrant, backing, qualifier* and *rebuttal*) in the students' posts. The presence of more components is held to be an indicator of higher argumentative quality. Analyzing the construction of single arguments, i.e., counting the presence of argument components, is referred to as the argument's micro-structure (Weinberger & Fischer, 2006).

A majority of researchers use a simplified version of Toulmin's model, which merges categories such as data with the warrant and backing under the heading *backing* and, then, the rebuttal and qualifier under the heading *qualification* (compare Figure 1 and Figure 2).

INSERT FIG:.1

Figure 2

Simplified Toulmin Model



Students' posts that lack *backing* or *qualification* are assumed to represent weaker argument quality than complete arguments.

Two aspects characterize how the included studies varied in terms of how Toulmin's model was employed in the analysis of online educational discussions: First, there was variation regarding the use of categories from the complete original model (Figure 1) or categories from a simplified Toulmin model (Figure 2). There was also variation concerning

whether the analysis focused solely on the microstructure of the arguments, i.e., the presence of the components of the Toulmin model, or whether other argument dimensions were measured, such as the argumentation's macro-structure, conceptual quality, etc. An overview of how the included studies were related to these two aspects is presented in the Appendix, Table 2, and summarized below. To illustrate the variation in how researchers use the Toulmin model for the analysis of argumentation in online educational discussions, the next sub-sections present examples of different coding schemes that employ categories from Toulmin's model. Following this, there is a summary of the frequency of the variation among the included studies according to whether a complete or simplified Toulmin model was used and whether the focus was multi-dimensional or exclusively in relation to the argument's micro-structure.

3.1. Coding schemes based on a simplified Toulmin model, supplemented with dimensions focusing on other aspects of argumentation

Weinberger and Fischer's (2006) coding scheme exemplifies a scheme that supplements coding of argument quality based on the categories of a simplified Toulmin model, in combination with additional aspects of argument quality. They suggest the following categories of the argument's micro-structure:

INSERT TABLE 2

Table 2

Weinberger & Fischer's (2006) Categories of Argument Micro-Structure Based on a Simplified Toulmin Model (Figure 2)

Category	Description	Component
Simple claim	Statement that raises a claim without providing any kind of evidence or limiting the context of where the claim is asserted to be valid and true	CLAIM
Grounded claim	Statement that raises a claim and provides evidence that proves the claim but lacks a limitation of the context wherein the claim is valid	CLAIM BACKING
Qualified claim	Statement that raises a claim and limits the context where it is true and valid but lacks evidence that proves its truth	CLAIM QUALIFICATION
Grounded and qualified claim	Statement that raises a claim, provides evidence that proves the truth and demarcates the context wherein the claim is valid	CLAIM BACKING QUALIFICATION
Non-argumentative moves	Posts that do not raise a substantial claim about the topic discussed, including questions, suggestions on how to proceed with the discussion, etc.	

Their scheme adds the following aspects of online educational discussions:

a) A *participation dimension*, focusing on both overall activity and the heterogeneity of interaction.

b) An *epistemic dimension*, focusing on the relation that discussants take toward the topic of discussion: the construction of the problem space, the conceptual space, the adequate relations between the problem and concepts, the inadequate relations between the problem and concepts, and the relations between prior knowledge and problem.

c) An *argument dimension*, which includes:

   1. A *micro-level*, focusing on how arguments are backed by evidence and limited. This dimension was built on the Toulmin model, as shown in Table 5.
   2. A *macro-level*, focusing on the function that a discussion post/single argument has in relation to other posts/arguments.

d) A *social mode of co-construction*, focusing on the function of a discussion post, for instance, articulating thoughts, questioning, coordinating the discussion, integration-oriented consensus building, and conflict-oriented consensus building.

3.2. Coding schemes based on the complete or simplified Toulmin model, without additional argument dimensions.

Cho and Jonassen (2002) used a coding scheme that strictly adapts Toulmin's argument components (complete model), coding for claims, grounds, warrants, backings, and rebuttals. The occurrence of each component is counted in each of the students' posts and scored on a scale from zero to six points, where complete arguments receive a higher score. This is based on the idea that the presence of a greater number of components represents a higher level of argumentative quality. Other researchers have used a simplified model that merged some categories; for instance, Liu et al. (2019) merged the grounds and warrant categories into a single category and coded each of the four remaining categories on a two-point scale.

3.3. Summed up

The starting point for my analysis was the assumption that Toulmin's argument model has a prominent status in research on online educational discussions (Weltzer-Ward, 2011; Wise & Paulus, 2016). The literature searches nuanced this picture, and an analysis of articles identified by Noroozi et al. (2012) revealed that alternative approaches to analyzing argumentation are as common as analyses of argumentation based on Toulmin's model. Furthermore, the analysis revealed variation in how researchers use the model. Nevertheless, all variations in the use of Toulmin's model take the presence of argument components as an indicator of argument quality, i.e., when more components are present in an argument, the argument quality is assumed to be higher.

INSERT TABLE 3

Table 3

Use of Different Coding Schemes Based on the Toulmin Model

	Uni-dimensional	**Multi-dimensional**
Complete Toulmin model	Complete Toulmin model, without additional argument dimensions **5 of 40 studies from Table 3** Example: Cho and Jonassen (2002)	
Simplified Toulmin model	Simplified Toulmin model, without additional argument dimensions **3 of 40 studies from Table 3** Example: Liu et al. (2019)	Simplified Toulmin model and dimensions focusing on other aspects of argumentation **32 of 40 studies from Table 3** Example: Weinberger and Fischer (2006)

Table 3 displays variation in how researchers have translated Toulmin's model into coding schemes for analyzing discussion transcripts. While some researchers employ all the argument components from the complete original model (Figure 1), others build on the simplified version of Toulmin's model (Figure 2), which is more commonly used than the former in analyzing students' arguments. Researchers use multi-dimensional coding, which combines categories from Toulmin's model with additional dimensions of argumentation (such as the "argumentation's macro-structure," "epistemic dimensions," and "conceptual quality"), more frequently than coding based solely on Toulmin's model. An important question is how variations in coding between the complete vs. the simplified Toulmin model and uni-dimensional vs. multi-dimensional coding influence the adequacy of these coding strategies. In the next section, I present and discuss caveats in using Toulmin's model as a measure of the quality of argumentation and how the dimensions identified in Table 3 (simplified vs. complete Toulmin model and uni-dimensional vs. multidimensional) relate to these caveats.

4. Discussion: How Adequate are the Frameworks that Build on the Toulmin Model for Analyzing Student Argumentation?

The second research question for this study concerns the adequacy of the Toulmin model as a basis for coding schemes used to analyze the quality of argumentation. Despite the

wide recognition of the model, scholars have criticized it as a general model of argumentation (Inch & Warnick, 2011; Nussbaum, 2011) and have questioned its use in measuring argumentation quality in online educational discussions (Andriessen & Baker, 2014; Noroozi et al., 2012). Remarkably, the researchers behind 14 of the 89 studies reviewed by Noroozi et al. (2012) discussed the Toulmin model and found it unsuitable for their analysis or argumentation in online educational discussions.

Two objections are central to these critiques: 1) The use of Toulmin's model for analysis focuses only on the proponent's side of an argumentative dispute and overlooks the contextual and dialogical aspects of argumentation. Argumentation is a dialogical activity in which one part—the proponent—tries to persuade another part about what to believe or do. This may occur for several reasons, for example, persuasion, inquiry (weighting arguments to decide what to believe), negotiation, information seeking, deliberation, and eristic argumentation (aimed at winning an argumentative conflict). Thus, argumentation always plays out in given contexts. In the next section, I discuss counting the presence of argument components in single arguments, disconnected from the argumentative context, as an approach to analyzing the quality of arguments. This caveat is most urgent when argumentation is analyzed using the Toulmin model alone, without focusing on additional dimensions of argumentation. 2) The second objection is identifying that several distinct components (claim, backing warrant, rebuttal, etc.) may be useful when explaining how arguments work and that it is cumbersome and not self-evidently fruitful when the model is used to measure the quality of arguments. The more fine-grained categories used, the more cumbersome and less reliable is the analysis. Thus, this objection is most urgent when analyzing arguments using the categories from the complete Toulmin model.

In the following sections, I will use arguments that are for and against the flat Earth theory as examples to demonstrate how one can be led astray when evaluating these arguments by counting the presence of Toulmin's argument components.

4.1. Counting Argument Components vs. Evaluating Relevance, Acceptability, and Sufficiency

Toulmin's model draws attention to the composition of single arguments and describes several argument components: claim, data, backing, warrant, qualifier, and rebuttal. As shown in previous sections, coding schemes based on the model take the presence of argument

components as an indicator of an argument's quality—more argument components equal a higher-quality argument. Is it reasonable to assess argument quality based on the presence of argument components? Let us consider an example:

> It is proven that a ship, which sails away from you toward the horizon, does not sink behind a hill of water but that it is actually perspective that hides it from view. This demonstrates that the Earth is not a globe. It has been found that the sinking ship effect is purely perceptual and that a good telescope with a sufficient zoom will change the observer's perspective and bring the ship's hull back into full view. This would not be possible if the ship were really behind a hill of water, since rays of light will not travel in a curved path. Hence, the effect that is usually thought to prove that the Earth is a globe really proves it to be a plane. (Adapted from The Flat Earth Society webpage[1].)

Coding this argument according to Toulmin's categories may look like:

- *Claim:* Earth is not a globe (i.e., the Earth is flat).
- *Data:* When you observe a ship that sails away from you toward the horizon, it does not sink behind a hill of water but is actually hidden by perspective.
- *Backing:* A good telescope with a sufficient zoom will change the observer's perspective and bring the ship's hull back into full view.
- *Warrant:* Rays of light will not travel in a curved path.
- *Rebuttal:* The sinking ship effect is usually thought to prove that the Earth is a globe.
- *Qualifier:* The 1proven (assuming that the conclusion must consequently be true).

Analyzed using the Toulmin model, this example contains a complete and, thus, high-quality argument. It demonstrates how the analysis of argument quality by counting the presence of argument components entails the risk of a false positive—an argument is taken to represent high quality, even when it is weak or even horrible, by other reasonable conceptions of argument quality.

Like Toulmin, several scholars have developed conceptions of argumentation that replace the deductive ideal, whereby conclusions necessarily follow from premises. In line with this, they have suggested criteria on which to judge the quality of a claim's justification.

[1] http://www.theflatearthsociety.org/tiki/tiki-index.php?page=Ships+appear+to+sink+as+they+recede+past+the+horizon (accessed 30.05.2019).

Johnson and Blair (2006, see also Blair, 2012) have suggested *relevance*, *acceptability*, and *sufficiency* as the criteria for accepting a claim's justification and, thus, for accepting claims. *Relevance* concerns the extent to which arguments or argument components support the conclusion they intend to support or whether they have a weak impact on the case. *Acceptability* concerns whether arguments or argument components are justified so that they can be accepted as true. *Sufficiency* concerns the weight of arguments and argument components—does the information presented provide enough justification for the acceptance of the overall claim? A line of reasoning such as the example above may be complete according to Toulmin's model but *irrelevant*, *unacceptable* (meaning poorly backed), or *insufficient* in a given argumentative context.

When categories from Toulmin's model are applied to the above example, there is no evaluation of the quality of the argument components or the quality of the relations between them. Alternatively, the arguments in the above example might be evaluated by the criteria of relevance, acceptability, and sufficiency within an argumentative context (e.g., a debate concerning the shape of the Earth). Problems with the above argument include the use of weak and dubious data, leading to a dubious claim; normally, when observing a ship disappear beyond the horizon, one will lose sight of the hull first and the mast last, even when using a telescope. Its warrant is also problematic—even if light normally travels in straight lines, atmospheric refraction may cause an optical illusion. Thus, the argument appears to lack acceptability, relevance, and sufficiency. An argument may score well with respect to the presence of the components of Toulmin's model and yet be horribly weak.

Analyzing argument quality by focusing on the composition of single arguments and the presence of argument components overestimates the significance of the presence of the components and underestimates alternative criteria, such as relevance, acceptability, and sufficiency. A more holistic approach to evaluating the quality of argumentation proposes that irrelevant arguments or arguments with poorly backed claims represent a poor quality of argumentation, even in the presence of all the Toulmin's argument components. In addition, coding the presence of the argument components alone, without considering the actual context, excludes the possibility of focusing on argumentative fallacies, including ad hominem and straw man arguments.

Some of the previously described studies employed multi-dimensional coding schemes. The problems discussed above affect such coding schemes to a lesser degree. For

instance, Clark and Sampson's (2008) coding scheme contains the dimensions of *discourse moves* (relating to an argument's relevance in an argumentative context) and *conceptual quality* (relating to the acceptability of an argument's evidence). Despite this, there is reason to question how the presence of argument components adds to the analysis of argument quality. For uni-dimensional coding schemes, which merely evaluate the quality of student arguments in terms of the presence of Toulmin-model components, there is a greater risk of missing out on the decisive qualities of arguments.

4.2. Challenges in Identifying Argument Components in Natural Language

The second objection to Toulmin's model as a basis for the coding schemes used to assess argumentative quality involves the challenges posed by identifying argument components in natural language (Erduran et al., 2004; Nussbaum, 2011). As stated, the key idea in such coding schemes is that the presence of argument components is a proper indicator of argument quality. Nevertheless, discussants do not normally formulate and construct their arguments according to the Toulmin model.[2] When discussants put forward their arguments, they regularly neglect to specify the elements they intend to use as their claim, data, evidence, etc. They also omit the argument components they deem to be obvious and, thus, implicit in the given context. Because of the sometimes implicit character of naturally occurring language, it is cumbersome to code student postings according to a coding scheme based on Toulmin's predefined conception of arguments: What should count as a claim, data, evidence, etc., are not necessarily evident.

Consider the following example:

> If you are able to watch a ship sail off to sea, watch its mast and flag as it fades off into the distance. You will notice that, in fact, it does not "fade off into the distance" at all; instead, you will see its mast and flag appear to slowly sink. The ship will have sailed beyond the point at which you would see it. Just to be sure, bring a pair of binoculars with you so that you can see even farther off into the distance. It's as if

[2] Some teaching and research designs use the Toulmin model as both a prescriptive model and a model to measure the quality of arguments. Students are instructed to form their utterances according to the model and to label parts of their arguments (see, e.g., Brooks & Jeong, 2006; A. Jeong, 2006; Jeong & Davidson-Shivers, 2006; Jeong & Frazier, 2008a, 2008b; Jeong & Joung, 2007; Jeong & Lee, 2008; Jeong, 2003, 2005; A. C. Jeong, 2006; Jeong, 2007). Nevertheless, objections toward the model as a means to analyze argument quality still apply.

> you're watching it go over to the other side of a hill. This phenomenon can only be explained by a sphere-shaped planet.[3]

Analyzing this excerpt, what would be reasonable to interpret as the main claim? Is the claim backed by evidence? Are there any warrants, rebuttals, and qualifiers? Even though most of us would agree with the claim in this argument—analyzed according to the coding schemes by focusing on the presence of the argument's components and disregarding the argumentative context—the argument seems to be implicit, incomplete, and even overconfident. There is nothing indicating what the claim in the argument actually is or what should count as justification. Coding this argument may go like this:

- *Claim:* This phenomenon (a ship that sails beyond the horizon seems to sink) can only be explained by a sphere-shaped planet.
- *Data:* No explicit data: "you will see its mast and flag appear to slowly sink" may count as data.
- *Backing:* Missing: there is no part of the argument that explicitly guarantees that the data function as evidence for the claim.
- *Warrant:* Missing: the argument has no component that supports its backing because there is no explicit backing in the argument. One might add that "light travels in straight lines" is the general principle that makes the observation possible (that a ship seems to sink when moving beyond the horizon) to be evidence for the claim.
- *Rebuttal:* There is no part of the argument that acknowledges possible counterarguments.
- *Qualifier:* "[C]an only be explained" may be interpreted as an estimate of the argument's probability. (Nevertheless, even if one accepts the argument, one might also accept that other explanations, such as an optical illusion, are possible.)

Evaluated solely on the basis of the presence of Toulmin's argument components, this argument would receive a low score. Knowing the context, a reasonable interpretation of the claim may be "the Earth is not flat" or "one of the most famous flat Earth arguments is

[3] Taken from "Crosstalk: 7 ways to prove the earth is round" http://crosstalk.cell.com/blog/seven-ways-to-prove-earth-is-round (accessed 30.05.2019).

wrong." When the context is known, coding seems to be easier. In argumentative interactions, discussants commonly omit argument components because we consider them implicit and, thus, unnecessary. The example illustrates that measuring argument quality by counting the presence of argument components may entail the risk of false negatives—an argument may be categorized as low quality, even when, in a given context, it holds an acceptable quality according to reasonable criteria for argument quality.

The coding schemes based on a simplified Toulmin model merge the six original components into three: the claim, backing, and qualifier (Figure 2), making the coding less cumbersome and avoiding the challenges posed by separating the data, backing, and warrant (or the qualifier and rebuttal). In the simplified Toulmin model, emphasis is placed on the overall structure of the arguments, not on the possible omission of components found to be redundant or implicit (typical warrants). Thus, the pitfalls of assessing argument quality by counting the presence of argument components appear to be more problematic when the more fine-grained argument component categories are applied.

5. Summary and Conclusion

Rational argumentation is a key competence in almost every academic discipline and is vital for deep learning and scrutinizing knowledge claims. Toulmin's argument model has been influential for studies of argumentation (Erduran et al., 2004; Leitao, 2000; Nussbaum, 2011; Weinberger & Fischer, 2006) and is held to be extensively employed in research on student argumentation in online educational discussions (Weltzer-Ward, 2011; Wise & Paulus, 2016). This article has provided an overview of how Toulmin's model is employed in the research field and discussed whether the model represents an adequate approach to the analysis of argumentative quality.

The literature search conducted for this current study demonstrated that Toulmin's model is commonly used in the research field, although alternative approaches to analyzing the quality of argumentation are equally important and frequently employed. Toulmin's model continues to be adapted in different ways. While some researchers adapt the model quite literally, counting all components from the original model (Figure 1), others merge components from the original model into broader categories (Figure 2). Some researchers use the presence (and/or absence) of Toulmin's argument components as the sole criterion for argumentative quality, while most researchers have added argument quality dimensions (e.g.,

conceptual quality) to the analysis of student postings. This article adds nuance to the perception that Toulmin's model is used extensively to conceptualize learning in online educational discussions. Further, the article describes variations in how the model is adapted for analysis.

Despite the influence of Toulmin's argument model, scholars have criticized its use as a method for analyzing argumentative quality (Andriessen & Baker, 2014; Inch & Warnick, 2011; Noroozi et al., 2012; Nussbaum, 2011). The presence of Toulmin's argument components does not necessarily imply high argumentative quality. Further, when focusing on the composition of single arguments (as Toulmin's model does), one may overlook the fact that arguments occur in contexts and that context matters when evaluating argument quality. Coding schemes that combine several argument dimensions (e.g., conceptual quality or discourse moves) in the analysis of student postings are, to a lesser degree, undermined by this critique. Nevertheless, an explanation of how the presence of argument components constitutes a sufficient, necessary, or even relevant dimension of argument quality is urgent. Further, scholars have remarked that naturally occurring language does not coincide with Toulmin's schematized analysis of single arguments. As a result, counting the presence of argument components in order to interpret and code natural language is cumbersome and risks disregarding some of the obvious reasons that discussants may have when omitting components in specific contexts. Identifying argument components is more challenging when fine-grained categories are employed. Thus, analyzing student arguments according to a simplified Toulmin model (Figure 2) is less problematic than using the more fine-grained categories derived from the complete model (Figure 1). However, analyses based on the merged categories from the simplified model still rely on the premise that counting the presence of argument components is a worthwhile approach to analyzing argument quality. While this critique is not entirely new, it has not been absorbed in the research field. Still, researchers publish studies that use this approach without considering the challenges described here.

Rejecting Toulmin's model as a means to analyze the quality of argumentation does not imply a rejection of Toulmin's contribution to the scholarly understanding of argumentation and critical thinking and how these relate to learning. Toulmin's work has been influential as an orientation toward informal logic, providing an alternative approach to the focus of classical logic on deduction and logical form. His approach highlights that providing

reasons (namely backing, data, ground, evidence, etc.) to consider a claim as acceptable is an essential feature of rational argumentation. While this does not imply that discussants should explicate all argument components every time they propose an argument (neither should they do this on a regular basis), being able to explicate upon request the components of one's argument strengthens the chance that one's argument will be evaluated as acceptable. The emphasis on the justification of claims as essential for rational argumentation coincides with a key element in most accounts of critical thinking. Despite confusion about how to define critical thinking,[4] most accounts of critical thinking acknowledge the evaluation of the tenability of claims as indispensable. Although Toulmin's model highlights this, he did not intend to present a normative model prescribing how good arguments should be composed (Nussbaum, 2011; Toulmin, 2003/1958, pp. vii-viii). Rather, the model served to analyze how the tenability of claims relates to their backing and how different argumentative contexts emphasize different standards for backing and warrant.

Previous reviews (De Wever et al., 2006; Noroozi et al., 2012) have mentioned the need to discuss how theoretical bases are operationalized for the analysis of online educational discussions. This article discussed how the Toulmin model is used in the research field as well as its adequacy as a means to analyze argumentative quality. Questions about the conceptualization of rational argumentation and critical thinking in education are crucial, not only as they relate to research on online educational discussions but also in more general debates about education, generic skills, and deep learning. Parts of the discussion raised may have value in related discussions concerning rational argumentation and critical thinking in other educational contexts, including the formulation of learning outcomes, teaching designs to enhance such competencies, and formative and summative assessments of such competencies.

[4] See, for instance, Siegel (2013, p. 2): "… the notion of critical thinking remains obscure and ill-defined; the theoretical conflicts between the various analyses offered to date are significant."

Appendices

Appendix, table 1. Criteria for Literature Search

Aim	Overview of studies using the Toulmin-model for content analysis of online educational discussions	
Databases	Education Research Complete (ERC) Education Research Information Center (ERIC)	
Limitations	• Articles published in peer reviewed, academic journals • English language	
Inclusion / exclusion criteria	Inclusion criteria: • Using Toulmin-model in analysis of online educational discussions, or • promoting one specific coding scheme for such analysis[5]	Exclusion criteria: • Using Toulmin-model for analysis of discussions or arguments in non-online settings[6] • Articles reviewing and comparing several coding schemes or frameworks, without promoting any one of them[7]

[5] See for instance the seminal article by Weinberger and Fischer (2006). This influential article presents a framework/coding scheme, yet no empirical analysis is included in the article.
[6] This excludes a large group of research that uses the Toulmin-model as an analytic tool in researching student argumentation in classroom settings; see for instance the works of (Erduran, 2018; Erduran et al., 2004).
[7] Like for instance De Wever et al. (2006), Clark et al. (2007) Spatariu, Hartley, and Bendixen (2004); Spatariu, Winsor, Simpson, and Hosman (2016).

Appendix, table 2. Articles using Toulmin's model for analysis of online, educational, discussions

	Author(s)	Title	Journal	Year	Search	Use of Toulmin's categories
1	Marttunen	Electronic mail as a pedagogical delivery system: An analysis of the learning of argumentation.	*Research in Higher Education*	1997	Noroozi & al. (2012)	Simplified Toulmin-model, and additional argument-dimensions
2	Cho & Jonassen	The effects of argumentation scaffolds on argumentation and problem solving	*Educational Technology Research and Development*	2002	Noroozi & al. (2012)	Complete Toulmin-model, without additional argument-dimensions
3	Mcalister, Ravenscroft & Scanlon	Combining interaction and context design to support collaborative argumentation using a tool for synchronous CMC.	*Journal of Computer Assisted Learning*	2004	Noroozi & al. (2012)	Simplified Toulmin-model, and additional argument-dimensions
4	Campos, Laferriere & Lapointe	Analysing Arguments in Networked Conversations: The Context of Student Teachers.	*Canadian Journal of Higher Education*	2005	Hand search	Simplified Toulmin-model, and additional argument-dimensions
5	Jeong	The Effects of Linguistic Qualifiers and Intensifiers on Group Interaction and Performance in Computer-Supported Collaborative Argumentation	*International Review of Research in Open and Distance Learning*	2005	Noroozi & al. (2012)	Simplified Toulmin-model, and additional argument-dimensions
6	Taasoobshirazi & Hickey	Promoting argumentative discourse: A design-based implementation and refinement of an astronomy multimedia curriculum, assessment model, and learning environment	*Astronomy Education Review*	2005	Noroozi & al. (2012)	Complete Toulmin-model, without additional argument-dimensions
7	Brooks & Jeong	Effects of pre-structuring discussion threads on group interaction and group performance in computer-supported collaborative argumentation	*Distance Education*	2006	Noroozi & al. (2012)	Simplified Toulmin-model, and additional argument-dimensions
8	Jeong	Gender interaction patterns and gender participation in computer-supported collaborative argumentation.	*The American Journal of Distance Education*	2006	Noroozi & al. (2012)	Simplified Toulmin-model, and additional argument-dimensions
9	Jeong	The effects of conversational language on group interaction and group performance in computer-supported collaborative argumentation	*Instructional Science*	2006	Noroozi & al. (2012)	Simplified Toulmin-model, and additional argument-dimensions

10	Jeong & Davidson-Shivers	The effects of gender interaction patterns on student participation in computer-supported collaborative argumentation	*Educational Technology Research and Development*	2006	Noroozi & al. (2012)	Simplified Toulmin-model, and additional argument-dimensions
11	Weinberger & Fischer	A framework to analyze argumentative knowledge construction in computer-supported collaborative learning	*Computers & Education*	2006	H	Simplified Toulmin-model, and additional argument-dimensions
12	Clark & Sampson	Personally-Seeded Discussions to Scaffold Online Argumentation	*International Journal of Science Education*	2007	Noroozi & al. (2012)	Simplified Toulmin-model, and additional argument-dimensions
13	Jeong	The effects of intellectual openness and gender on critical thinking processes in computer-supported collaborative argumentation	*Journal of Distance Education,*	2007	Noroozi & al. (2012)	Simplified Toulmin-model, and additional argument-dimensions
14	Jeong & Joung	Scaffolding collaborative argumentation in asynchronous discussions with message constraints and message labels	*Computers & Education*	2007	Noroozi & al. (2012)	Simplified Toulmin-model, and additional argument-dimensions
15	Kollar, Fischer & Slotta,	Internal and external scripts in computer-supported collaborative inquiry learning	*Learning and Instruction*	2007	Noroozi & al. (2012)	Simplified Toulmin-model, and additional argument-dimensions
16	Stegmann, Weinberger, & Fischer	Facilitating argumentative knowledge construction with computer-supported collaboration scripts	*International Journal of Computer-Supported Collaborative Learning*	2007	Noroozi & al. (2012)	Simplified Toulmin-model, and additional argument-dimensions
17	Clark & Sampson	Assessing Dialogic Argumentation in Online Environments to Relate Structure, Grounds, and Conceptual Quality	*Journal of Research in Science Teaching*	2008	Noroozi & al. (2012)	Simplified Toulmin-model, and additional argument-dimensions
18	Cross, Taasoobshirazi, Hendricks & Hickey	Argumentation: A strategy for improving achievement and revealing scientific identities	*International Journal of Science Education*	2008	Noroozi & al. (2012)	Simplified Toulmin-model, and additional argument-dimensions
19	Golanics & Nussbaum	Enhancing online collaborative argumentation through question elaboration and goal	*Journal of Computer Assisted Learning*	2008	Noroozi & al. (2012)	Simplified Toulmin-model, and additional argument-dimensions
20	Jeong & Frazier	How Day of Posting Affects Level of Critical Discourse in Asynchronous Discussions and Computer-Supported Collaborative Argumentation	*British Journal of Educational Technology*	2008	Noroozi & al. (2012)	Simplified Toulmin-model, and additional argument-dimensions
21	Jeong & Lee	The Effects of Active versus Reflective Learning Style on the Processes of Critical Discourse in	*British Journal of Educational Technology*	2008	Noroozi & al. (2012)	Simplified Toulmin-model, and additional

		Computer-Supported Collaborative Argumentation				argument-dimensions
22	Clark, D'Angelo & Menekse	Initial Structuring of Online Discussions to Improve Learning and Argumentation: Incorporating Students' Own Explanations as Seed Comments versus an Augmented-Preset Approach to Seeding Discussions	*Journal of Science Education and Technology*	2009	Noroozi & al. (2012)	Simplified Toulmin-model, and additional argument-dimensions*
23	Huang, Wang, Huang, Chen, Chen & Chang	Performance Evaluation of an Online Argumentation Learning Assistance Agent	*Computers & Education*	2011	Hand search	Complete Toulmin-model, without additional argument-dimensions
24	Lu, Chiu, & Law	Collaborative argumentation and justifications: A statistical discourse analysis of online discussions	*Computers in Human Behavior*	2011	Hand search	Simplified Toulmin-model, and additional argument-dimensions
25	Asterhan, Schwarz & Gil	Small-group, computer-mediated argumentation in middle-school classrooms: The effects of gender and different types of online teacher guidance.	*British Journal of Educational Psychology*	2012	Hand search	Simplified Toulmin-model, and additional argument-dimensions
26	Lin, Hong, & Lawrenz	Promoting and scaffolding argumentation through reflective asynchronous discussions.	*Computers & Education*	2012	Hand search	Simplified Toulmin-model, and additional argument-dimensions
27	Stegmann, Wecker, Weinberger & Fischer	Collaborative argumentation and cognitive elaboration in a computer-supported collaborative learning environment	*Instructional Science*	2012	Hand search	Simplified Toulmin-model, and additional argument-dimensions
28	Noroozi, Weinberger, Biemans, Mulder & Chizari	Facilitating argumentative knowledge construction through a transactive discussion script in CSCL.	*Computers & Education*	2013	Hand search	Simplified Toulmin-model, and additional argument-dimensions
29	Choi, Hand & Norton-Meier	Grade 5 Students' Online Argumentation about Their In-Class Inquiry Investigations	*Research in Science Education*	2014	Hand search	Simplified Toulmin-model, and additional argument-dimensions
30	Tsai & Tsai	College students' skills of online argumentation: The role of scaffolding and their conceptions.	*Internet & Higher Education*	2014	Hand search	Complete Toulmin-model, without additional argument-dimensions
31	Wise, Hausknecht & Zhao	Attending to others' posts in asynchronous discussions: Learners' online "listening" and its relationship to speaking.	*International Journal of Computer-Supported Collaborative Learning,*	2014	Hand search	Simplified Toulmin-model, and additional argument-dimensions
32	Özçinar	Scaffolding Computer-Mediated Discussion to Enhance Moral Reasoning and Argumentation Quality in Pre-Service Teachers	*Journal of Moral Education*	2015	Hand search	Complete Toulmin-model, without additional argument-dimensions

33	Camus, Hurt, Larson & Prevost	Facebook as an Online Teaching Tool: Effects on Student Participation, Learning, and Overall Course Performance	*College Teaching*	2016	Hand search	Simplified Toulmin-model, and additional argument-dimensions*
34	Oh & Kim	Understanding Cognitive Engagement in Online Discussion: Use of a Scaffolded, Audio-Based Argumentation Activity	*International Review of Research in Open and Distributed Learning*	2016	Hand search	Simplified Toulmin-model, and additional argument-dimensions
35	Weng, Lin & She	Scaffolding for Argumentation in Hypothetical and Theoretical Biology Concepts	*International Journal of Science Education*	2017	Hand search	Simplified Toulmin-model, without additional argument-dimensions
36	Chen, Chang, Ouyang & Zhou	Fostering student engagement in online discussion through social learning analytics	*Internet & Higher Education*	2018	Search 2	Simplified Toulmin-model, and additional argument-dimensions*
37	Lam, Hew & Chiu	Improving Argumentative Writing: Effects of a Blended Learning Approach and Gamification	*Language Learning & Technology*	2018	Search 2	Simplified Toulmin-model, and additional argument-dimensions
38	Wise & Hsiao	Self-regulation in online discussions: Aligning data streams to investigate relationships between speaking, listening, and task conditions	*Computers in Human Behavior*	2018	Search 2	Simplified Toulmin-model, and additional argument-dimensions
39	Lin	Student positions and web-based argumentation with the support of the six thinking hats	*Computers & Education*	2019	Search 2	Simplified Toulmin-model, without additional argument-dimensions
40	Liu, Liu & Lin	The influence of prior knowledge and collaborative online learning environment on students' argumentation in descriptive and theoretical scientific concept	*International Journal of Science Education*	2019	Search 2	Simplified Toulmin-model, without additional argument-dimensions
	See table 1, appendix table, and article text for search procedures and inclusion/exclusion criteria.					* indicates that there is no explicit reference to Toulmin in the article

References

Andriessen, J., & Baker, M. (2014). Arguing to Learn. In R. K. Sawyer (Ed.), *The Cambridge Handbook of the Learning Sciences* (pp. 439-460). Cambridge: Cambridge University Press.

Asterhan, C. S. C., Schwarz, B. B., & Gil, J. (2012). Small-Group, Computer-Mediated Argumentation in Middle-School Classrooms: The Effects of Gender and Different Types of Online Teacher Guidance. *British Journal of Educational Psychology, 82*(3), 375-397. Retrieved from. doi:10.1111/j.2044-8279.2011.02030.x

Barik, T., Ford, D., Murphy-Hill, E., & Parnin, C. (2018). *How Should Compilers Explain Problems to Developers?* Paper presented at the ESEC/FSE 2018 - Proceedings of the 2018 26th ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering.

Biasutti, M. (2017). A Comparative Analysis of Forums and Wikis as Tools for Online Collaborative Learning. *Computers & Education, 111*, 158-171. Retrieved from doi:https://doi.org/10.1016/j.compedu.2017.04.006

Blair, J. A. (2012). Relevance, Acceptability and Sufficiency Today. In C. W. Tindale (Ed.), *Groundwork in the Theory of Argumentation: Selected Papers of J. Anthony Blair* (pp. 87-100). Dordrecht: Springer Netherlands.

Brooks, C. D., & Jeong, A. (2006). Effects of Pre-Structuring Discussion Threads on Group Interaction and Group Performance in Computer-Supported Collaborative Argumentation. *Distance Education, 27*(3), 371-390.

Campos, M. N., Laferriere, T., & Lapointe, J. M. (2005). Analysing Arguments in Networked Conversations: The Context of Student Teachers. *Canadian Journal of Higher Education, 35*(4), 55-84. Retrieved from http://umanitoba.ca/outreach/csshe/Pub/pub.html.

Camus, M., Hurt, N. E., Larson, L. R., & Prevost, L. (2016). Facebook as an Online Teaching Tool: Effects on Student Participation, Learning, and Overall Course Performance. (Cover Story). *College Teaching, 64*(2), 84-94. Retrieved from. doi:10.1080/87567555.2015.1099093

Chen, B., Chang, Y.-H., Ouyang, F., & Zhou, W. (2018). Fostering Student Engagement in Online Discussion through Social Learning Analytics. *Internet & Higher Education, 37*, 21-30. Retrieved from doi:10.1016/j.iheduc.2017.12.002

Cho, K. L., & Jonassen, D. H. (2002). The Effects of Argumentation Scaffolds on Argumentation and Problem Solving. *Etr&D-Educational Technology Research and Development, 50*(3), 5-22. Retrieved from <Go to ISI>://WOS:000178942000001. doi:Doi 10.1007/Bf02505022

Choi, A., Hand, B., & Norton-Meier, L. (2014). Grade 5 Students' Online Argumentation About Their in-Class Inquiry Investigations. *Research in Science Education, 44*(2), 267-287. Retrieved from http://dx.doi.org/10.1007/s11165-013-9384-8.

Clark, D., Sampson, V., Weinberger, A., & Erkens, G. (2007). Analytic Frameworks for Assessing Dialogic Argumentation in Online Learning Environments. *Educational Psychology Review, 19*(3), 343-374. Retrieved from http://dx.doi.org/10.1007/s10648-007-9050-7. doi:10.1007/s10648-007-9050-7

Clark, D. B., D'Angelo, C. M., & Menekse, M. (2009). Initial Structuring of Online Discussions to Improve Learning and Argumentation: Incorporating Students' Own Explanations as Seed Comments Versus an Augmented-Preset Approach to Seeding Discussions. *Journal of Science Education and Technology, 18*(4), 321-333. Retrieved from http://dx.doi.org/10.1007/s10956-009-9159-1.

Clark, D. B., & Sampson, V. D. (2007). Personally-Seeded Discussions to Scaffold Online Argumentation. *International Journal of Science Education, 29*(3), 253-277.

Clark, D. B., & Sampson, V. D. (2008). Assessing Dialogic Argumentation in Online Environments to Relate Structure, Grounds, and Conceptual Quality. *Journal of Research in Science Teaching, 45*(3), 293-321. doi:10.1002/tea.20216

De Wever, B., Schellens, T., Valcke, M., & Van Keer, H. (2006). Content Analysis Schemes to Analyze Transcripts of Online Asynchronous Discussion Groups: A Review. *Computers & Education, 46*(1), 6-28. doi:http://dx.doi.org/10.1016/j.compedu.2005.04.005

De Wever, B., Van Keer, H., Schellens, T., & Valcke, M. (2007). Applying Multilevel Modelling to Content Analysis Data: Methodological Issues in the Study of Role Assignment in Asynchronous Discussion Groups. *Learning and Instruction, 17*(4), 436-447.

Duschl, R. A. (2007). Quality Argumentation and Epistemic Criteria. In *Argumentation in Science Education* (pp. 159-175): Springer.

Erduran, S. (2018). Toulmin's Argument Pattern as a "Horizon of Possibilities" in the Study of Argumentation in Science Education. *Cultural Studies of Science Education*, 1-9. doi:10.1007/s11422-017-9847-8

Erduran, S., Ozdem, Y., & Park, J.-Y. (2015). Research Trends on Argumentation in Science Education: A Journal Content Analysis from 1998–2014. *International Journal of STEM Education, 2*(1), 5. doi:10.1186/s40594-015-0020-1

Erduran, S., Simon, S., & Osborne, J. (2004). Tapping into Argumentation: Developments in the Application of Toulmin's Argument Pattern for Studying Science Discourse. *Science Education, 88*(6), 915-933. doi:10.1002/sce.20012

Gao, F., Zhang, T., & Franklin, T. (2013). Designing Asynchronous Online Discussion Environments: Recent Progress and Possible Future Directions. *British Journal of Educational Technology, 44*(3), 469-483.

Golanics, J. D., & Nussbaum, E. M. (2008). Enhancing Online Collaborative Argumentation through Question Elaboration and Goal Instructions. *Journal of Computer Assisted Learning, 24*(3), 167-180.

Huang, C.-J., Wang, Y.-W., Huang, T.-H., Chen, Y.-C., Chen, H.-M., & Chang, S.-C. (2011). Performance Evaluation of an Online Argumentation Learning Assistance Agent. *Computers & Education, 57*(1), 1270-1280.http://dx.doi.org/10.1016/j.compedu.2011.01.013.

Inch, E. S., & Warnick, B. (2011). *Critical Thinking and Communication: The Use of Reason in Argument*: Allyn & Bacon.

Janssen, J., Erkens, G., Kirschner, P. A., & Kanselaar, G. (2010). Effects of Representational Guidance During Computer-Supported Collaborative Learning. *Instructional Science, 38*(1), 59-88.

Jeong, A. (2006). Gender Interaction Patterns and Gender Participation in Computer-Supported Collaborative Argumentation. *The American Journal of Distance Education, 20*(4), 195-210.

Jeong, A., & Davidson-Shivers, G. V. (2006). The Effects of Gender Interaction Patterns on Student Participation in Computer-Supported Collaborative Argumentation. *Educational Technology Research and Development, 54*(6), 543-568.

Jeong, A., & Frazier, S. (2008a). How Day of Posting Affects Level of Critical Discourse in Asynchronous Discussions and Computer-Supported Collaborative Argumentation. *British Journal of Educational Technology, 39*(5), 875-887.

Jeong, A., & Frazier, S. (2008b). How Day of Posting Affects Level of Critical Discourse in Asynchronous Discussions and Computer-Supported Collaborative Argumentation. *British Journal of Educational Technology, 39*(5), 875-887.

Jeong, A., & Joung, S. (2007). Scaffolding Collaborative Argumentation in Asynchronous Discussions with Message Constraints and Message Labels. *Computers & Education, 48*(3), 427-445.doi:http://dx.doi.org/10.1016/j.compedu.2005.02.002

Jeong, A., & Lee, J. (2008). The Effects of Active Versus Reflective Learning Style on the Processes of Critical Discourse in Computer-Supported Collaborative Argumentation. *British Journal of*

Educational Technology, 39(4), 651-665. http://dx.doi.org/10.1111/j.1467-8535.2007.00762.x.

Jeong, A. C. (2003). The Sequential Analysis of Group Interaction and Critical Thinking in Online. *American Journal of distance education, 17*(1), 25-43. Retrieved from http://dx.doi.org/10.1207/S15389286AJDE1701_3. doi:10.1207/s15389286ajde1701_3

Jeong, A. C. (2005). The Effects of Linguistic Qualifiers and Intensifiers on Group Interaction and Performance in Computer-Supported Collaborative Argumentation. *International Review of Research in Open and Distance Learning, 6*(3). Retrieved from

Jeong, A. C. (2006). The Effects of Conversational Language on Group Interaction and Group Performance in Computer-Supported Collaborative Argumentation. *Instructional Science, 34*(5), 367.

Jeong, A. C. (2007). The Effects of Intellectual Openness and Gender on Critical Thinking Processes in Computer-Supported Collaborative Argumentation. *Journal of Distance Education, 22*(1), 1-18.

Johnson, R. H., & Blair, J. A. (2006). *Logical Self-Defense*: Idea.

Lam, Y. W., Hew, K. F., & Chiu, K. F. (2018). Improving Argumentative Writing: Effects of a Blended Learning Approach and Gamification. *Language Learning & Technology, 22*(1), 97-118. Retrieved from

Leitao, S. (2000). The Potential of Argument in Knowledge Building. *Human Development, 43*(6), 332-360.

Lin, H.-s., Hong, Z.-R., & Lawrenz, F. (2012). Promoting and Scaffolding Argumentation through Reflective Asynchronous Discussions. *Computers & Education, 59*(2), 378-384. doi:https://doi.org/10.1016/j.compedu.2012.01.019

Lin, Y.-R. (2019). Student Positions and Web-Based Argumentation with the Support of the Six Thinking Hats. *Computers & Education, 139*, 191-206. doi:https://doi.org/10.1016/j.compedu.2019.05.013

Liu, Q.-T., Liu, B.-W., & Lin, Y.-R. (2019). The Influence of Prior Knowledge and Collaborative Online Learning Environment on Students' Argumentation in Descriptive and Theoretical Scientific Concept. *International Journal of Science Education, 41*(2), 165-187. Retrieved from https://doi.org/10.1080/09500693.2018.1545100. doi:10.1080/09500693.2018.1545100

Lu, J., Chiu, M. M., & Law, N. W. (2011). Collaborative Argumentation and Justifications: A Statistical Discourse Analysis of Online Discussions. *Computers in Human Behavior, 27*(2), 946-955. doi:https://doi.org/10.1016/j.chb.2010.11.021

Marttunen, M. (1997). Electronic Mail as a Pedagogical Delivery System: An Analysis of the Learning of Argumentation. *Research in Higher Education, 38*(3), 345-363.

Maxwell, J. A. (2013). *Qualitative Research Design : An Interactive Approach* (3rd ed. ed. Vol. 41). Los Angeles: Sage.

Mcalister, S., Ravenscroft, A., & Scanlon, E. (2004). Combining Interaction and Context Design to Support Collaborative Argumentation Using a Tool for Synchronous Cmc. *Journal of Computer Assisted Learning, 20*(3), 194-204.

Noroozi, O., Weinberger, A., Biemans, H. J. A., Mulder, M., & Chizari, M. (2012). Argumentation-Based Computer Supported Collaborative Learning (Abcscl): A Synthesis of 15 Years of Research. *Educational Research Review, 7*(2), 79-106.doi:http://dx.doi.org/10.1016/j.edurev.2011.11.006

Noroozi, O., Weinberger, A., Biemans, H. J. A., Mulder, M., & Chizari, M. (2013). Facilitating Argumentative Knowledge Construction through a Transactive Discussion Script in Cscl. *Computers & Education, 61*(0), 59-76. doi:http://dx.doi.org/10.1016/j.compedu.2012.08.013

Nussbaum, E. M. (2011). Argumentation, Dialogue Theory, and Probability Modeling: Alternative Frameworks for Argumentation Research in Education. *Educational Psychologist, 46*(2), 84-106.

Oh, E. G., & Kim, H. S. (2016). Understanding Cognitive Engagement in Online Discussion: Use of a Scaffolded, Audio-Based Argumentation Activity. *International Review of Research in Open and Distributed Learning, 17*(5), 28-48..

Siegel, H. (2013). *Educating Reason*: Routledge.

Spatariu, A., Hartley, K., & Bendixen, L. D. (2004). Defining and Measuring Quality in Online Discussions. *The Journal of Interactive Online Learning, 2*(4), 1-15.

Spatariu, A., Winsor, D. L., Simpson, C., & Hosman, E. (2016). Further Classification and Methodological Considerations of Evaluations for Online Discussion in Instructional Settings. *Turkish Online Journal of Educational Technology, 15*(1), 43.

Stegmann, K., Wecker, C., Weinberger, A., & Fischer, F. (2012). Collaborative Argumentation and Cognitive Elaboration in a Computer-Supported Collaborative Learning Environment. *Instructional Science, 40*(2), 297-323. doi:10.1007/s11251-011-9174-5

Stegmann, K., Weinberger, A., & Fischer, F. (2007). Facilitating Argumentative Knowledge Construction with Computer-Supported Collaboration Scripts. *International Journal of Computer-Supported Collaborative Learning, 2*(4), 421-447. Retrieved from http://dx.doi.org/10.1007/s11412-007-9028-y. doi:10.1007/s11412-007-9028-y

Taasoobshirazi, G., & Hickey, D. T. (2005). Promoting Argumentative Discourse: A Design-Based Implementation and Refinement of an Astronomy Multimedia Curriculum, Assessment Model, and Learning Environment.

Toulmin, S. (2003/1958). *The Uses of Argument*. Cambridge: Cambridge University Press.

Tsai, P.-S., & Tsai, C.-C. (2014). College Students' Skills of Online Argumentation: The Role of Scaffolding and Their Conceptions. *The Internet and Higher Education, 21*, 1-8. doi:http://dx.doi.org/10.1016/j.iheduc.2013.10.005

Weinberger, A., & Fischer, F. (2006). A Framework to Analyze Argumentative Knowledge Construction in Computer-Supported Collaborative Learning. *Computers & Education, 46*(1), 71-95. doi:http://dx.doi.org/10.1016/j.compedu.2005.04.003

Weltzer-Ward, L. (2011). Content Analysis Coding Schemes for Online Asynchronous Discussion. *Campus-Wide Information Systems, 28*(1), 56-74.

Weng, W.-Y., Lin, Y.-R., & She, H.-C. (2017). Scaffolding for Argumentation in Hypothetical and Theoretical Biology Concepts. *International Journal of Science Education, 39*(7), 877-897. http://dx.doi.org/10.1080/09500693.2017.1310409.

Wise, A. F., Hausknecht, S. N., & Zhao, Y. (2014). Attending to Others' Posts in Asynchronous Discussions: Learners' Online "Listening" and Its Relationship to Speaking. *International Journal of Computer-Supported Collaborative Learning, 9*(2), 185-209. Retrieved from https://doi.org/10.1007/s11412-014-9192-9. doi:10.1007/s11412-014-9192-9

Wise, A. F., & Hsiao, Y.-T. (2018). Self-Regulation in Online Discussions: Aligning Data Streams to Investigate Relationships between Speaking, Listening, and Task Conditions. *Computers in Human Behavior*. doi:10.1016/j.chb.2018.01.034

Wise, A. F., & Paulus, T. M. (2016). Analyzing Learning in Online Discussions. *The SAGE Handbook of E-learning Research*, 270.

Özçinar, H. (2015). Scaffolding Computer-Mediated Discussion to Enhance Moral Reasoning and Argumentation Quality in Pre-Service Teachers. *Journal of Moral Education, 44*(2), 232-251. http://dx.doi.org/10.1080/03057240.2015.1043875.