JOURNAL OF BALTIC SCIENCE EDUCATION

ISSN 1648-3898 /Print/
ISSN 2538-7138 /Online/

TRENDS AND FOUNDATIONS IN RESEARCH ON STUDENTS' CONCEPTUAL UNDERSTANDING IN SCIENCE EDUCATION: A METHOD BASED ON THE STRUCTURAL TOPIC MODEL

Shuaishuai Mi,
Shanshan Lu,
Hualin Bi

Abstract. *This study aims to reveal the trends and foundations in research on students' conceptual understanding in science education. The literature was selected from three recognized journals in science education. The Structural Topic Model (STM) method was used to categorize articles into ten topics considering information about the semantic cohesion and exclusivity of words to topics. The topic, which has attracted increasing research interest, was selected using a method similar to standard regression analysis, and its changing focus was identified through an analysis of its research contents. Foundations of research about students' conceptual understanding between 1980–1999 and 2000–2019 were obtained through a review of their top 10 most-cited papers. Three conclusions were drawn: a) there were ten sub-topics of research about students' conceptual understanding; b) the research on the development (or pathways) of students' scientific argumentation/reasoning is likely to attract further interest in the future; and c) compared to the studies in the first period, the studies in the second stage favor research on the description (nature, mental process, etc.) of the process of students' conceptual understanding as the research foundation.*

Keywords: *conceptual understanding, journal publication, structural topic model, text mining.*

Shuaishuai Mi,
Shanshan Lu,
Hualin Bi
Shandong Normal University, China

Introduction

Concepts and principles are the basic building blocks of scientific knowledge, and understanding a concept is a prerequisite to making complex inferences or accomplishing any scientific work with it (Reif, 1995). Therefore, it is crucial for students to attain conceptual understanding. Enormous interest from science educators has been captured by research into students' conceptual understanding, and many articles on it have been published since 1980 (Tsai & Wen, 2005). However, it is not clear what aspects of students' conceptual understanding these papers were concerned with, or which aspects are likely to attract more interest from researchers. This information might be important for researchers to reflect upon within their own research foci and plan further investigations (Lin et al., 2019). Therefore, this study focused on the trends and foundations of research on students' conceptual understanding.

Before identifying the trends and foundations of present research, it is necessary to define students' conceptual understanding. The research on students' conceptual understanding was initiated by the research on the constructivist view of learning, knowledge, and understanding. According to the constructivist view of learning, the growth of understanding always involves a learner constructing his or her own private understanding of some part of the public knowledge. Here, students' own private understanding was often described as misconception (the use of "mis-" applies only when one implicitly compares a student's private conception with some public, accepted meaning (Bezzi, 1996; Pines & West, 1986)). Of course, students' private understanding can be described not only as misconceptions (Brumby, 1984; Cho et al., 1985; Gallegos et al., 1994; Taber, 2014) but also as preconceptions, alternative conceptions, alternative frameworks, alternative conceptual frameworks, intuitive theories, mini-theories (Taber, 2015a), prior knowledge (Taber, 2015b), and students' conceptions (a collection of related and interrelated conceptions that students hold, see M. Hewson & P. Hewson, 1983).

To illustrate the above constructing process, Posner et al. (1982) used ***conceptual change*** to describe it. Conceptual change means a people's central, organizing concept changing from one set of concepts to a different set that is incompatible with the first (Posner et al., 1982). To date, conceptual

https://doi.org/10.33225/jbse/20.19.551

ISSN 1648–3898 /Print/
ISSN 2538–7138 /Online/

change has been one of the research strands of the National Association for Research in Science Teaching (NARST) conference (Lin et al., 2019).

The terms used to describe the constructing process also included conceptual development (a process of coming to view one theory or model as having more explanatory power than others, see Taber, 2001; or the production of coherent mental maps, see Gilbert et al., 2011) and learning progression (mapping individual students' performance to reveal a picture of conceptual change in the domain over time, see Claesgens et al., 2002; Salinas, 2009).

In sum, in this study, students' conceptual understanding is regarded as a process of transformation of students' experiences from the initial state (private understanding) to the final state (some part of the public knowledge). Thus, the research on students' conceptual understanding includes the research on students' private understanding (described by misconception, alternative conception, etc.), and the research on the constructing process (described by conceptual change, conceptual development, etc.).

Literature Review

Since this study concerns the trends and foundations of research on students' conceptual understanding, the trends should be discussed first. What, then, do the trends of research on students' conceptual understanding mean? Trends are often regarded as the interest in a research topic (Cuccurullo et al., 2016; Dukes & Strauch, 1984), and with the help of these, researchers can reflect on their own research foci and plan further investigations (Lin et al., 2019). Therefore, some studies have analyzed the trends of science education, and some features of research on students' conceptual understanding in science education has been included in these papers.

White (1997) gathered counts of keywords from summaries of articles in the Educational Research Information Clearinghouse (ERIC). Then, he divided the years from 1966 to 1995 into six five-year periods. Finally, the paper presented the number of times the keywords appeared per 10,000 articles in each period. White found that the instances of conceptions or misconceptions grew from 0 to 355 in these six five-year periods. The results indicated that the research on conceptions or misconceptions captured the most interest from researchers from 1965 to 1996.

The research trends of science education in the period 1998–2017 are reported in four articles: Tsai and Chang (2005), Lee et al. (2009), Lin et al. (2014), and Lin et al. (2019). The methods used in these four papers were similar. Specifically, the literature for analysis was selected from three widely recognized journals: specifically, International Journal of Science Education (IJSE), Journal of Research in Science Teaching (JRST), and Science Education (SE). Then, authors referred the criteria from the National Association for the Research in Science Teaching (NARST) conference categories (https://narst.org/conferences) and classified literature (excluding "editorials," "commentaries," "responses," and "book reviews") into only one artificial category. Finally, the trends of nine topics were summarized through the comparison of percentages of the quantity of publications on them. Tsai and Chang (2005) analyzed the trends in science education from 1998 to 2002. They concluded that, although the research topic about Learning-Conception was the most frequently investigated one, a declining trend was observed when analyzed by year. Lee et al. (2009) analyzed the papers published in all three journals from 2003 to 2007. Similar to the years 1998 to 2002, Learning-Conceptions revealed a declining trend. In contrast to 1998–2007 (Lee et al., 2009; Tsai & Chang, 2005), the declining trend of Learning-Conceptions decelerated in the period from 2008–2012 (Lin et al., 2014). In the period of 2013–2017, Learning-Conceptions was still one of the top research topics (Lin et al., 2019), although the percentage of publications on the research topic of Learning-Conceptions fell lower than the period, 2008–2012. In sum, these four papers indicated that though Learning-Conceptions has been one of the top research topics, it has been in rapid decline over the last 20 years. A similar result was obtained by Jong (2007) after analyzing the articles published from 1995 to 2005 in IJSE, JRST, and SE. Chang et al. (2010) analyzed the papers published from 1990 to 2007 in the four journals of IJSE, JRST, SE and Research in Science Education (RISE), and nine topics were obtained, using complete linkage clustering. This research presented a similar result: although conceptual change and concept mapping attracted most studies in the field of science education, a declining trend was observed after 2000.

Two results regarding research on students' conceptual understanding can be obtained from the above-mentioned studies. One is that research on students' conceptual understanding has been a focus of science education research (Lee et al., 2009; Lin et al., 2014; Lin et al., 2019; Tsai & Wen, 2005). The second result is that, although the research on students' conceptual understanding has been one of the top research topics in science education, the publications on it have shown an upward trend from 1965 to 2000s (Chang et al., 2010; White, 1997) and a downward trend was revealed after 2000 (Chang et al., 2010; Lee et al., 2009; Lin et al., 2014; Lin et al., 2019; Tsai &

https://doi.org/10.33225/jbse/20.19.551

Wen, 2005). It seems contradictory that this research area has been one of the top three research foci of science education for almost twenty years, but with a downward trend of publications for this same period. How can we account for this confusing conclusion? In the present study, two possible reasons were considered.

One reason for this phenomenon may be the popularity of research on students' conceptual understanding to the varied sub-topics within it (Lin et al., 2019). It can be explained that, as an area of research, there are some research contents that have been evoked, and simultaneously, there are some research contents that have been vanishing. Therefore, an upward trend for some sub-topics may be accompanied by a downward trend for others. The downward trend of the research on conceptual understanding might be attributed to the process of change in the prevalence of sub-topics. For example, the number of sub-topics with upward trends might be smaller than the number with downward trends, and in terms of the quantity of publications of research on students' conceptual understanding, it might show a downward trend. In fact, the research on students' conceptual understanding covers many sub-topics, including the investigation of students' misconception (or alternative conception, or alternative framework, see Banerjee, 1991 and Simpson & Marek, 1988) and the instructional interventions to help students overcome misconceptions or to facilitate students' conceptual change (Basili & Sanford, 1991; Stavy, 1991). Unfortunately, no article reported sub-topics of research on students' conceptual understanding in the above. As the information about sub-topics was not provided, the upward (or downward) trends of sub-topics cannot be identified, nor can the change of interest in research on students' conceptual understanding. Therefore, the first purpose of this study is to identify the sub-topics of research on students' conceptual understanding.

The second reason for the abnormal trends of the research may be attributed to the variety of methods used to obtain trends. White (1997) obtained trends through summarizing the number of times the keywords appeared per 10,000 articles in all time periods. Tsai and Wen (2005), Lee et al. (2009), Lin et al. (2014), and Lin et al. (2019) evaluated trends through the comparison of percentages of quantity of publications on them. Chang et al. (2010) obtained trends through a line chart of the quantity of articles belonging to a specific topic published annually. Though the above-mentioned studies used different methods to obtain the trends of research on students' conceptual understanding, they all used descriptive statistical methods to deduct the research trends. However, such descriptive statistical methods cannot be used to make a statistical prediction (Aron et al., 2013). Therefore, even if the information on the sub-topics of research on students' conceptual understanding is obtained, the change in interest from researchers cannot be predicted using the methods mentioned above. Therefore, the second issue in the present study was to obtain the trends of sub-topics in research on students' conceptual understanding using a method that could make a statistical prediction.

To address the above issues, the first step is to obtain the sub-topics of research on students' conceptual understanding. There are two ways to do this according to the previous studies. The first method is to use the criteria from the NARST conference categories to analyze the sub-topics (Lee et al., 2009; Lin et al., 2014; Lin et al., 2019; Tsai and Wen, 2005). There are two reasons this method did not work well: no sub-topics of students' conceptual understanding reported in the criteria from NARST conference categories and the inevitably subjective assignment of articles (Chang et al., 2010). The second method is provided by Chang et al. (2010) (complete linkage clustering), and this method is also designed to reduce the subjectivity of assignment. This method calculates the similarity between articles, then employs a similarity threshold to combine different articles into clusters. This process repeats until no clusters can be merged. Finally, each of the articles can be assigned to a cluster. Simultaneously, the sub-topics are obtained. One fact to be noticed here is that all the articles belong to research on students' conceptual understanding. Therefore, when the complete linkage clustering with the similarity threshold 0.05 is used to obtain the topics, most articles would be assigned to one topic (before this study was conducted, we tried this method and the results showed that it did not work well). Therefore, a different method needed to be used to obtain the sub-topics. This method was capable of obtaining the sub-topics of research on students' conceptual understanding in a non-preset style (reducing the subjective assignment of articles). In addition to being able to avoid the subjective assignment of articles, this method was also able to obtain the trends of sub-topics using a method that can make a statistical prediction.

The Structural Topic Model (STM) is a model used for dividing collections of documents into natural groups (Roberts et al., 2014a). It divides articles into different groups using a method for unsupervised classification (Chandelier et al., 2018), which is different from supervised methods, where the topics are defined by hand-coding a corpus of documents (Banks et al., 2018). This method infers topics from the texts rather than assuming them prior to analysis (Banks et al., 2018), which means the sub-topics are obtained in a non-preset style. Therefore, STM can not only divide documents into groups based on texts but also eliminate much of the subjectivity of ratings and

https://doi.org/10.33225/jbse/20.19.551

ISSN 1648–3898 /Print/
ISSN 2538–7138 /Online/

categorizations. The STM has additional advantages that make it as a proper model to analyze the research trends about students' conceptual understanding. The special advantages of STM include allowing researchers to estimate the relationship (in a way that is similar to standard regression analysis) between topics and document metadata (such as publication date, see Chandelier et al., 2018; Roberts et al., 2014a), which is a characteristic that the other topic modeling models (such as Latent Dirichlet Allocation, see Blei et al., 2003 etc.) do not possess. Simultaneously, this is necessary to address the second issue mentioned above. Therefore, STM was used to obtain the trends of the research on students' conceptual understanding in this study. This study is not the first to use STM to obtain trends of research topics; it has been used in political (Dybowski & Adämmer, 2018; Geese, 2019; Shirokanova & Silyutina, 2018) and education sciences (Reich et al., 2015) (the published research that used STM can be found at www.structuraltopicmodel.com).

This study is concerned not only with the trends of research on students' conceptual understanding but also the foundations of this research. In every field of science, select publications play an important role in the development of the field. Certain articles become an expediting factor in the development of the field because of their influence (Berry & Parasuraman, 1993), which is the basis of the present study. That is, the studies that play an important role in the development of research on students' conceptual understanding are regarded as its foundations. In the above-mentioned studies, highly cited references in science education were analyzed by the researchers (Lee et al., 2009; Lin et al., 2014; Lin et al., 2019; Tsai & Wen, 2005). However, the foundations of the research on students' conceptual understanding have not been concerned with such studies. Therefore, the third objective of this study is to obtain the foundations of the research on students' conceptual understanding in science education by summarizing the 10 most cited papers.

Research Questions

The above analysis revealed three significant questions to be answered:

1. What sub-topics are included in the research on students' conceptual understanding?
2. How did the sub-topics of the research on students' conceptual understanding vary with time?
3. How did the 10 most highly cited studies of research on students' conceptual understanding vary with time?

Research Methodology

Data Collection and Pre-processing

The Web of Science database contains more than 10,000 world authoritative and high-impact journals, covering the fields of physical sciences, social sciences, technology, the earliest of which dates back to 1900 (Zhang et al., 2015). For many years, Web of Science has had a virtual monopoly on the provision of citation tracking (Bakkalbasi et al., 2006). Therefore, Web of Science has often been used as a source of data.

The journals widely used as sources for selecting articles in science education included International Journal of Science Education (IJSE), Journal of Research in Science Teaching (JRST), and Science Education (SE) (Lee et al., 2009; Lin et al., 2014; Lin et al., 2019; Tsai & Chang, 2005). Chang et al. (2010) also used the journal of Research in Science Education (RISE) in addition to the three recognized journals above due to similarly high impact and zonal characteristics (JRST, SE in North America, IJSE, in Europe, and RISE in Australasia).

In the present study, IJSE, JRST, SE, and RISE were all used as sources for selecting articles in consideration of impact and zonal characteristics. Thus, the publication name (SO) was set as SO= ("SCIENCE EDUCATION" OR "INTERNATIONAL JOURNAL OF SCIENCE EDUCATION" OR "JOURNAL OF RESEARCH IN SCIENCE TEACHING" OR "RESEARCH IN SCIENCE EDUCATION") in Web of Science.

As the definition of students' conceptual understanding states, students' private understanding was described using several terms (e.g., misconception, alternative framework), and the constructing process was described by several terms (e.g., conceptual change, conceptual development). These terms were grouped as a suite of terms to represent the research on students' conceptual understanding. Thus, the topic was set as TS=("misconcept*" OR "preconception*" OR "alternat* concept*" OR "alternat* framework*" OR "alternat* concept* framework*" OR "intuit* theor*" OR "mini-theor*" OR "prior knowledge*" OR "student* concept*" OR "concept* change" OR "concept* develop*" OR "learn* progress*")(quotation marks indicated, for example, alternat* framework* were combined as a word group, * was used to indicate synonyms of a term, for example, misconception instead of misconceptions).

https://doi.org/10.33225/jbse/20.19.551

In addition, the language types were confined to English, and the document types were confined to articles. As the introduction section mentioned that numerous studies on students' conceptual understanding have been published since the 1980s (Lin et al., 2019), only the studies published in or after 1980 were selected and the time span was set as "1980 to 2019" in this study. The date of data collection was February 12, 2020. Finally, the results showed that a total of 1,388 articles were suitable according to the searching criteria.

An abstract is a brief and comprehensive summary of the contents of an article that can correctly reflect the purpose and content of a manuscript (APA, 2009). Readers can grasp the necessary information of an entire paper through the abstract without reading all the words or sentences in the article (Zhang et al., 2015). In this study, a corpus of article abstracts was used to obtain the sub-topics of research about students' conceptual understanding. Therefore, the 33 studies without abstracts were excluded.

It was necessary to engage in some processing of the textual data before modeling it. The most common processing steps included lemmatization (Benoit & Matsuo, 2019), dropping punctuation, dropping numbers, and stopping word removal (e.g., the, is, at) (Roberts et al., 2014a). In this study, the copyright information presented in some abstracts was also removed.

Statistical Analysis

The STM package was used to conduct the statistical analysis according to the description of Roberts et al. (2014a) about usage (Roberts et al., 2018), namely model search and understanding. With regard to model search, the most important step was to select an appropriate number of topics. There is no right answer to the number of appropriate topics for a given corpus (Grimmer & Stewart, 2013). Based on the recommendations of Roberts et al. (2014b), the topic quality was measured through words that were semantically cohesive and exclusive to topics. Thus, 10 sub-topics were regarded as appropriate.

Understanding means the interpretation of model results. The first aspect of understanding is to comprehend topics through words and example articles. Chandelier et al. (2018) described the selected topics based on the inspection of the 20 top words and the reading of the 5 abstracts that were most representative of each topic. In the present article, we referred in the same manner to defined sub-topics mentioned in research on students' conceptual understandings in science education. The second aspect of understanding is to estimate relationships between metadata and topical content (Roberts et al., 2014a). As the literature review section stated, trends refer to the change of interest in a topic over time. Therefore, the publication time was set as a covariate to obtain the change of interest in sub-topics. As for the research interest in a topic, there are typically two characteristics of concern: the prevalence and the research content (Roberts et al., 2014b). In sum, the second aspect of understanding is to estimate relationships between publication time of a study and the prevalence and research content. In addition to the above, to obtain the trends of topics efficiently, the topic correlations were estimated using STM.

As mentioned in the literature review section, there was an upward trend in the publications of research on students' conceptual understanding from 1965 to 2000 (Chang et al., 2010; White, 1997). The year 2000 was a defining period in the publications of this research. After 2000, there was a downward trend (Chang et al., 2010; Lee et al., 2009; Lin et al., 2014; Lin et al., 2019; Tsai & Wen, 2005). Simultaneously, only articles published in or after 1980 were included in this study. Therefore, in this study, two periods of time were examined: 1980–1999 and 2000–2019. Therefore, the change in research contents was obtained through the comparison of the most frequent words between the two periods. Furthermore, foundations of research on students' conceptual understanding within the two periods were derived through an analysis of the 10 most highly cited papers in the two periods.

Research Results

20 Most Probable Words of Topics and 5 Papers Most Highly Associated with Each Topic

The 20 most probable words of each topic were as follows:

Topic 1: Hypothetico, evolutionary theory, evolutionary explanation, species, leverage, evolution, explanatory hypothesis, evolutionary change, natural selection, loss, survival, combustion, adaptation, trait, biological evolution, worksheet, predictive, Greek, item, progression.

Topic 2: Intuitive rule, teacher belief, pedagogical knowledge, elementary school child, initial teacher

https://doi.org/10.33225/jbse/20.19.551

ISSN 1648–3898 /Print/
ISSN 2538–7138 /Online/

training, main category, pedagogical content knowledge (PCK), vignette, content, teacher education, content knowledge, science method, teacher view, physicist, teaching practice, physics teacher, opinion, constructivism, professional development.

Topic 3: Philippines, cellular respiration, subsequent, Science-Technology (ST), epistemological view, assertion, pre-service teacher view, scientific community, science teacher education program, ambiguity, intelligible, science method, teaching approach, expectation, simultaneous, prospective teacher, classroom discourse, ST instruction, meet, classroom activity.

Topic 4: Low achiever, dependent variable, teacher demonstration, anchor, student achievement, achiever, boy, gender, covariance, work, prior knowledge, diagnostic test, achievement, repeat, visualization, electricity, girl, factor analysis, sufficiently, treatment.

Topic 5: Gender difference, future study, incline, survey, genetic, depict, adoption, science learning, relationship, school student, male, positively, meaningful, negatively, semester, deep, questionnaire, predict, female, worldview.

Topic 6: Analogy, analog, metaphor, analogically, analogical reasoning, exert, bridge analogy, visit, student teacher, literacy, bridge, situational, analogue, classroom discourse, table, teacher educator, root, education level, clinical, light.

Topic 7: Refutational text, control student, Lakatos, plot, experimental treatment, exercise, heat, experimental, treatment, log, achievement test, teach experiment, cooperative, text, Cambridge, promising, simulation, concept mapping, quasi, curiosity.

Topic 8: Velocity, ozone layer, boiling, impose, condensation, relativity, greenhouse effect, reconstruct, presupposition, anode, atmosphere, air pressure, solute, aqueous solution, naive view, character, pupil, existence, distance, pre-instructional.

Topic 9: Science teaching efficacy, psychometric, operation, diagnosis, Le Chatelier, reasoning ability, analogical model, bond, gas behavior, electron, diagnose, invertebrate, respiration, freshman, item difficulty, alternative conceptual framework, tier, vertebrate, meiosis.

Topic 10: Reinterpret, rejection, anomalous datum, induce, taxonomy, reject, uncertainty, cognitive conflict, response, exclude, accept, respond, scientific conception, 9th, stimulate, ignore, final, logical thinking ability, socio-scientific, conflict map.

To understand the research content of each topic, Roberts et al. (2014a) suggested that both the collections of words associated with topics and the actual documents estimated to be highly associated with each topic be examined. Thus, for each topic, five documents were selected by examining the abstracts, which were highly associated with each topic, as shown in Table 1.

Table 1
Top 5 Documents that are highly associated with each topic

Topic	Document 1	Document 2	Document 3	Document 4	Document 5
Topic 1	LAWSON, AE, 1992, 10.1002/ TEA.3660290205	PEEL, A, 2019, 10.1080/09500693.2018.1564084	ZANGORI, L, 2017, 10.1002/ TEA.21404	ROMINE, WL, 2017, 10.1002/ TEA.21380	PEEL, A, 2019, 10.1002/ TEA.21545
Topic 2	GROSSSCHEDL, J, 2014, 10.1080/09500693.2014.923949	COLLINS, A, 1993, 10.1002/ TEA.3660300908	PITJENG-MOSABALA, P, 2018, 10.1080/09500693.2018.1446569	STAVY, R, 1996, 10.1080/0950069960180602	TIROSH, D, 1996, 10.1080/0950069960180603
Topic 3	AKERSON, VL, 2000, 10.1002/ (SICI)1098-2736(200004)37:4<295::AID-TEA2>3.3.CO;2-U	TABACHNICK, BR, 1999, 10.1002/(SICI)1098-237X(199905)83:3<309::AID-SCE3>3.0.CO;2-1	MESCI, G, 2017, 10.1007/S11165-015-9503-9	HANUSCIN, DL, 2006, 10.1002/SCE.20149	TSAI, CC, 1999, 10.1080/095006999290156
Topic 4	TRUMPER, R, 2001, 10.1080/09500690010025085	WU, HC, 2010, 10.1007/S11165-009-9138-9	STAVER, JR, 1988, 10.1002/ TEA.3660250906	YEH, TK, 2012, 10.1080/09500693.2011.579640	CHAMBERS, SK, 1997, 10.1002/(SICI)1098-2736(199702)34:2<107::AID-TEA2>3.0.CO;2-X

https://doi.org/10.33225/jbse/20.19.551

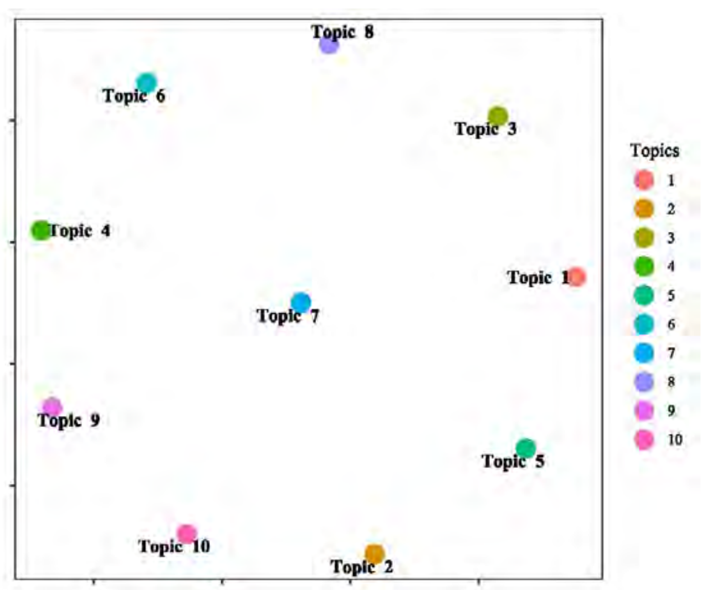
ISSN 1648–3898 /Print/
ISSN 2538–7138 /Online/

Topic	Document 1	Document 2	Document 3	Document 4	Document 5
Topic 5	LIN, TJ, 2014, 10.1080/09500693.2013.780317	CAVALLO, AML, 1994, 10.1002/TEA.3660310408	SCHMIDT, HJ, 1995, 10.1080/0950069950170605	CHIOU, GL, 2012, 10.1080/09500693.2011.558131	LIN, TC, 2015, 10.1080/09500693.2014.992057
Topic 6	BROWN, DE, 1994, 10.1080/0950069940160208	NIEBERT, K, 2012, 10.1002/SCE.21026	AFONSO, AS, 2007, 10.1002/SCE.20220	BROWN, BA, 2016, 10.1002/SCE.21212	LAWSON, DI, 1993, 10.1002/TEA.3660301012
Topic 7	ZACHARIA, ZC, 2008, 10.1002/TEA.20260	OLYMPIOU, G, 2012, 10.1002/SCE.20463	RYOO, K, 2014, 10.1002/TEA.21128	ZHANG, ZH, 2013, 10.1080/09500693.2013.792971	NIAZ, M, 1995, 10.1002/TEA.3660320907
Topic 8	FELZMANN, D, 2014, 10.1080/09500693.2014.936328	JOUNG, YJ, 2009, 10.1080/09500690701744603	EBENEZER, JV, 2001, 10.1002/SCE.1021	BOYES, E, 1993, 10.1080/0950069930150507	GRECA, IM, 2002, 10.1002/SCE.10013
Topic 9	NEUMANN, K, 2013, 10.1002/TEA.21061	ABRAHAM, MR, 1994, 10.1002/TEA.3660310206	CALEON, IS, 2010, 10.1007/S11165-009-9122-4	ROMINE, WL, 2016, 10.1002/SCE.21240	OTHMAN, J, 2008, 10.1080/09500690701459897
Topic 10	LIN JY, 2007, 10.1002/TEA.20125	CHINN, CA, 1998, 10.1002/(SICI)1098-2736(199808)35:6<623::AID-TEA3>3.0.CO;2-O	KANG, SJ, 2004, 10.1023/B:RISE.0000021001.77568.B3	ZEIDLER, DL, 2002, 10.1002/SCE.10025	CHEN, YC, 2019, 10.1002/SCE.21527

The Correlations between Topics

The correlations between topics reveal how topics correlate with one another through co-occurrence at the document level (Roberts et al., 2014a). This assists in getting a sense of how likely a single document discusses any given set of topics (Bohr and Dunlap, 2018). In this study, as mentioned in the research methodology section, the information on correlations of topics has provided additional evidence for the inference pertaining to the prevalence of topics. Figure 1 shows the correlations between topics.

Figure 1
The correlations between topics



As shown in Figure 1, ten topics are divided into ten clusters. There is no correlation between the topics. This result indicated that there are no correlations between the changes in quantity of articles belonging to different topics.

https://doi.org/10.33225/jbse/20.19.551

ISSN 1648–3898 /Print/
ISSN 2538–7138 /Online/

The Prevalence of Sub-Topics

As mentioned in the research methodology section, the change in prevalence of topics over time was accounted for. STM can do this in a manner that is similar to standard regression analysis, where topic-proportions are the outcome variable (Roberts et al., 2018). Table 2 lists the results of the estimations of regressions.

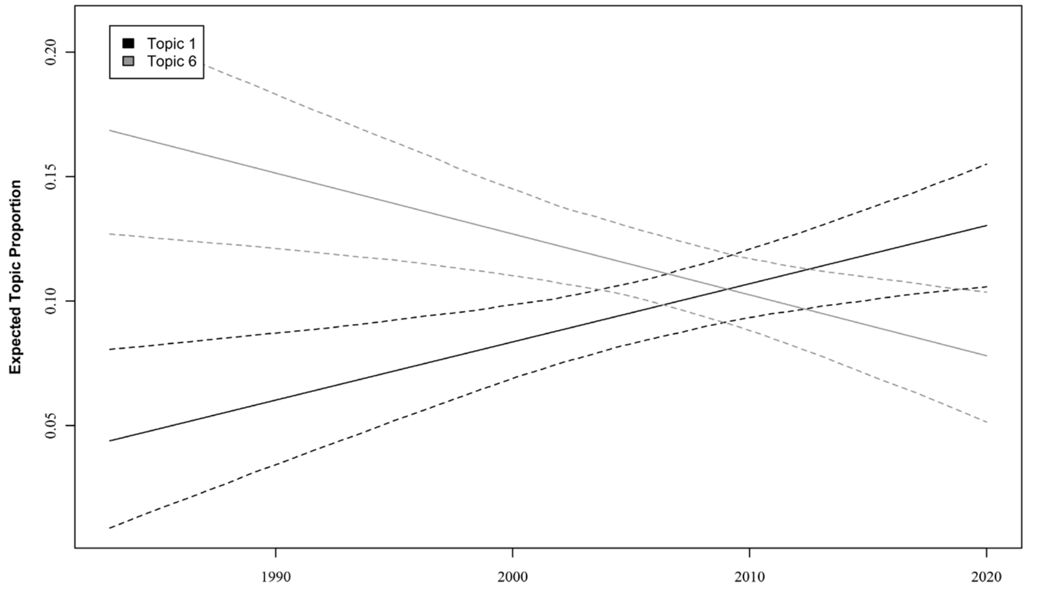
Table 2
The prevalence of topics

Topic	Estimate	*t* value	*p* (>|t|)
1	.002263	2.603035	.009341**
2	.000693	.927234	.353971
3	-.000624	-.809862	.418162
4	.000517	.737848	.460735
5	.000814	1.281784	.200138
6	-.002327	-2.997701	.002770*
7	.000796	.982515	.326022
8	-.001480	-1.734005	.083145
9	-.000567	-.712903	.476028
10	-.000085	-.287854	.773502

*Note. **p<.01; *p<.05; estimates=regression coefficient*

As Table 2 shows, the results indicate that there is a statistical significance of topic 1 (t=2.60, p<.01) and topic 6 (t=-2.99, p<.05). Figure 2 shows the change in the prevalence of topic 1 and topic 6 over the years. As shown in Figure 2, the proportion of topic 1 shows an upward trend from 1980, and the proportion of topic 6 shows a downward trend from 1980.

Figure 2
The prevalence of topic 1 and topic 6.

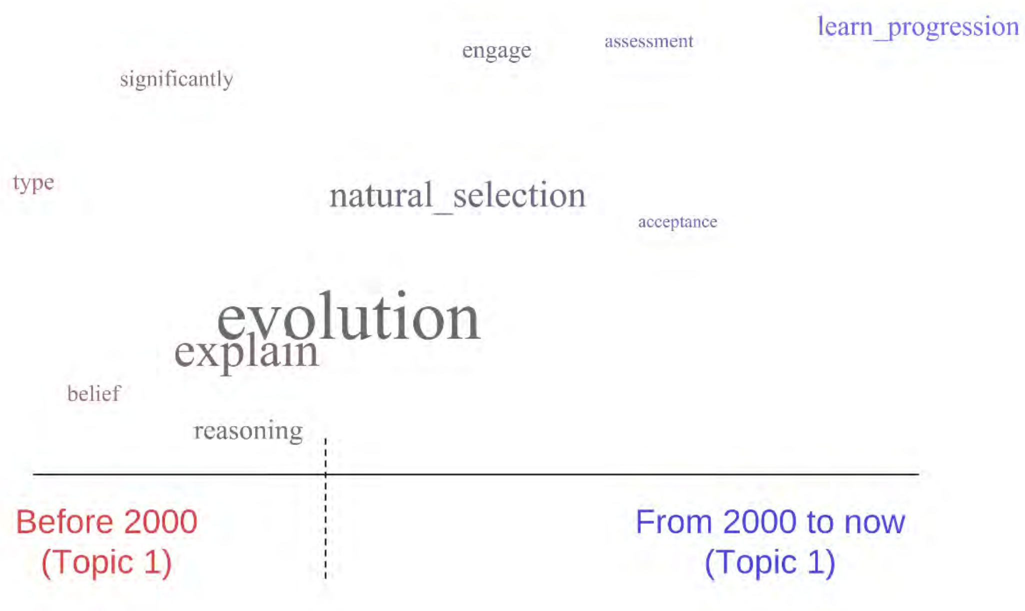


https://doi.org/10.33225/jbse/20.19.551

The Change of Research Contents in Sub-Topics over Time

Besides the change in prevalence over time, the change in research topics also includes the change in research content. As the above section stated, only the regression coefficient (estimate) of topic 1 reached a significant level at 0.05 with a positive value. Thus, only the change in research contents in topic 1 is presented in this section. As mentioned in the method section, the years from *1980 to 2019* are divided into two periods, before and after *2000*. Therefore, the change in research contents in topic 1 is shown in these two periods, as shown in Figure 3.

Figure 3
The change in research contents in topic 1



Note. Words are sized proportionally to their use within the topic-covariate combinations and oriented along the X-axis based on how much they favor one of the two configurations.

As Figure 3 shows, the phrases *type*, *belief*, *significantly*, *reasoning*, and *explain* are exclusive to the period before *2000*, while the phrases *natural selection, engage, assessment, acceptance,* and *learning progression* are exclusive to the period after *2000*. Evolution is shared with the content of the two periods, which indicates the focus of the research on evolution across the two periods.

The 10 Most Highly Cited Papers in the Two Periods

As mentioned in the literature review section, highly cited papers were recognized as being of crucial importance to the development of research on students' conceptual understanding. By reviewing highly cited papers, the foundations of the research can be obtained. Table 3 shows the 10 most highly cited papers in the two periods, one from 1980 to 1999 and another from 2000 to 2019.

https://doi.org/10.33225/jbse/20.19.551

ISSN 1648–3898 /Print/
ISSN 2538–7138 /Online/

Table 3
The change in the 10 most highly cited papers in the two periods from 1980 to 2019

Before 2000	From 2000 to now
GILBERT, J. K., 1983, STUDIES SCI ED, V10, P61, DOI 10.1080/03057268308559905	POSNER, G.J, 1992, PHILOSOPHY OF SCIENCE, COGNITIVE PSYCHOLOGY, AND EDUCATIONAL THEORY AND PRACTICE, P147
DRIVER, R, 1983, STUDIES SCI ED, V10, P37, DOI 10.1080/03057268308559904	CAREY, S., 1985, CONCEPTUAL CHANGE IN CHILDHOOD
HEWSON, PW, 1989, INT J SCI EDUC, V11, P541, DOI 10.1080/0950069890110506	DRIVER, R., 1994, MAKING SENSE OF SECONDARY SCIENCE
PINTRICH, PR, 1993, REVIEW OF EDUCATIONAL RESEARCH, V63, P167, DOI 10.3102/00346543063002167	VOSNIADOU, S., 1994, LEARN INSTR, V4, P45, DOI 10.1016/0959-4752(94)90018-3
GILBERT, J. K., 1982, SCI EDUC, V66, P623, DOI 10.1002/SCE.3730660412	DUIT, R, 2003, INT J SCI EDUC, V25, P671, DOI 10.1080/09500690305016
NOVAK, J.D., 1984, LEARNING LEARN	VOSNIADOU, S, 1992, COGNITIVE PSYCHOLOGY, V24, P535, DOI 10.1016/0010-0285(92)90018-W
FREYBERG, P, 1985, LEARNING IN SCIENCE. THE IMPLICATIONS OF CHILDREN'S SCIENCE	PINTRICH, PR, 1993, REV EDUC RES, V63, P167, DOI 10.3102/00346543063002167
DRIVER, R, 1985, CHILDRENS IDEAS SCI	DISESSA, AA, 1993, COGNITION INSTRUCT, V10, P105, DOI 10.1080/07370008.1985.9649008
DRIVER, R, 1978, STUDIES SCI ED, V5, P61, DOI 10.1080/03057267808559857	NATIONAL RESEARCH COUNCIL, 1996, NATIONAL SCIENCE EDUCATION STANDARDS
POSNER, GJ, 1982, SCI EDUC, V66, P211, DOI 10.1002/SCE.3730660207	POSNER, GJ, 1982, SCI EDUC, V66, P211, DOI 10.1002/SCE.3730660207

As Table 3 shows, there are two identical papers between the first period (from 1980 to 1999) and the second period (from 2000 to 2019) (Pintrich et al., 1993; Posner et al., 1982).

Discussion

Main Research Topics

The results section provided information about the 20 most probable topic phrases and the top five papers highly associated with each topic. With the results of these two types of information on each topic, the focus of topics was obtained. The process of obtaining the focus of topic 1 is shown below, using these two types of information.

Within the framework of STM, a topic is defined as a mixture of words where each word has a probability of belonging to a topic, which means the sum word probabilities for a given topic is one (Roberts et al., 2014a). Therefore, the 20 most probable words of topics presented in the results section are, in fact, the words with the highest probability of belonging to a given topic. As for topic 1, the collection of words that are associated with it include hypothetico, evolutionary theory, evolutionary explanation, specie, leverage, evolution, explanatory hypothesis, evolutionary change, natural selection, loss, survival, combustion, adaptation, trait, biological evolution, worksheet, predictive, Greek, item, progression. Among the above 20 phrases, hypothetico had the highest probability, which means it is the most representative phrase of topic 1. Hypothetico is one of the elements of hypothetico-predictive argumentation (or hypothetico-deductive), which represents a pattern of scientific argumentation (Lawson, 1985; Lawson, 2003). Lawson (2003), Lawson (1985), and Niaz (1988) presumed that the ability to construct and comprehend hypothetico-predictive (or hypothetico-deductive) arguments is necessary for the construction of conceptual knowledge due to its essential functions in concept construction and conceptual change. Given that explanatory hypothesis (the process of creating possible, alternative explanations for a given set of information, see Park, 2006) and predictive are included in the collection, it is suggested that topic 1 concerns some patterns of students' scientific argumentation, which play an important role in students'

https://doi.org/10.33225/jbse/20.19.551

conceptual understanding. As with the importance of scientific argumentation in students' conceptual understanding, the research of scientific argumentation has vast appeal to researchers. Moreover, it seems negligent to infer the focus of topics is using just three phrases (hypothetico, explanatory hypothesis, and predictive) since evolutionary theory, evolutionary explanation, specie, leverage, etc. are also included in the collection. Most phrases are related to scientific concepts except hypothetico, explanatory hypothesis, and predictive, which reflect the goal of students' conceptual understanding (or the goal can be defined as constructing students' own private understanding, see introduction section). These concepts can be discussed in relation to the research on students' scientific argumentation. For example, Lawson (2003) took Darwin's Theory of Evolution as an example to illustrate how hypothetico-predictive argumentation is involved in concept construction and conceptual change. Lavoie (1999) investigated the effect of emphasizing hypothetico-predictive reasoning within the science learning cycle on high school students' conceptual understandings in biology (natural selection), etc.

The 20 most probable phrases of topic 1 were provided by the STM package (Roberts et al., 2014b), which means all these phrases are highly associated with topic 1, and in this study, this meant that these phrases collectively represented the research about scientific argumentation. One fact that should be noticed here is that there is a possibility that these scientific concepts were discussed together in a topic that was not concerned with students' scientific argumentation, because these scientific concepts are the goal of students' conceptual understanding (these scientific concepts may have been discussed together in reference to a topic that concerned the investigation of students' conceptions). Therefore, though there is a reason to believe these twenty words strongly correlated with each other and were combined together to discuss students' scientific argumentation due to belonging to the same topic, it is premature to conclude that topic 1 concerned students' scientific argumentation only using the above twenty phrases.

To check the above judgments of the focus of topic 1, it is necessary to use the five studies highly associated with topic 1, which are presented in Table 1. The first study, conducted by Lawson and Worsnop (1992), found that the acquisition of domain-specific concepts largely depends on reflective reasoning skill. The second study, conducted by Peel et al. (2019b), investigated students' natural selection and antibiotic resistance (ABR) knowledge through model-based explanations. The third study, conducted by Zangori et al. (2017), theorized that embedding model-based reasoning within a socio-scientific issue (SSI) unit would support students in developing and using models to articulate explanations of carbon cycling, climate change, and their interrelationships. The fourth study, conducted by Romine et al. (2017), validated the Measure of Acceptance of the Theory of Evolution (MATE) on undergraduate students using the Rasch model and a path to mediate evolution acceptance might be reasoning and understanding (Ha et al., 2012). The fifth study conducted, by Peel et al. (2019a), analyzed students' pre- and post-unit algorithmic explanations of natural selection and obtained the change in students' conceptions of natural selection. It can be found that the studies highly associated with topic 1 were all concerned with students' scientific reasoning rather than with scientific argumentation, which was concluded from the phrases that are associated with topic 1. As Osborne (2010) and Fischer et al. (2014) stated, educational and science education research on scientific argumentation has focused on the externalized processes and products of scientific reasoning within social contexts. Therefore, there is no disagreement between the conclusion obtained from the phrases that are associated with topic 1 and the studies highly associated with topic 1.

In sum, by combining the information provided by the 20 most probable phrases of topic 1 and studies highly associated with topic 1, it can be concluded that topic 1 focuses on the research about scientific argumentation/reasoning in students' conceptual understanding. The foci of other topics are obtained through the same process mentioned above, as follows:

Topic 1: Scientific argumentation/reasoning, which plays an important role in the construction of students' conceptions.

Topic 2: Teachers' pedagogical content knowledge (content knowledge, pedagogical knowledge, etc.), which contributes to students' conceptual understanding.

Topic 3: Nature of science (both teachers' and students'), which can facilitate students' conceptual understanding.

Topic 4: The investigation of students' conceptions (or prior knowledge), and the relationships (e.g., mediators) between it and other factors (e.g., gender, treatment).

https://doi.org/10.33225/jbse/20.19.551

ISSN 1648–3898 /Print/
ISSN 2538–7138 /Online/

Topic 5: Conceptions of science learning (both teachers' and students'), which is of benefit to students' conceptual understanding.

Topic 6: Analogies and metaphors, which can enhance students' conceptual understanding.

Topic 7: Cognitive conflict, which is effective in promoting students' conceptual understanding.

Topic 8: Contexts, which have influence on the construction of students' specific conception.

Topic 9: The development of the diagnostic instruments of students' conceptions, and the source of the variation of students' conceptual understanding.

Topic 10: Students' response to the discrepant events, which can promote students' conceptual understanding.

The focus of each sub-topic is typically stated in a short statement. The purpose of obtaining the sub-topics of research on students' conceptual understanding is a sub-step to obtain the trends of research on students' conceptual understanding. Therefore, only the sub-topics that have been attracting more and more interest from researchers were used to develop and provide a full and detailed account of topics that might attract more interest from researchers.

It must be reemphasized that the foci of sub-topics were obtained by combining the information provided by the 20 most probable phrases of each topic and five documents that were highly associated with each topic, rather than simply using one or the other. This led to the result that the foci of some sub-topics seemed to conflict with the information provided by the 20 most probable phrases. For example, cognitive conflict was one of the 20 most probable phrases of topic 10, but the focus of topic 7, rather than the focus of topic 10, was described as cognitive conflict that can be effective in promoting students' conceptual understanding. In fact, this is attributed to the information provided by the five documents that were most highly associated with each topic. The five documents that were most highly associated with topic 7 focused on intervention to promote students' conceptual understanding; these interventions facilitated students' conceptual understanding through invoking a cognitive conflict. They include the instructional interventions of physical and virtual manipulations (Olympiou & Zacharia, 2012), and teaching strategies based on cognitive conflict (Niaz, 1995). The five documents that were most highly associated with topic 10 primarily focused on the students' responses to discrepant events (Kang et al., 2004) or anomalous data (Lin, J-Y, 2007). Therefore, in this study, the focus of topic 7, rather than the focus of topic 10, was identified as cognitive conflict, which can be effective in promoting students' conceptual understanding. This followed the recommendations of Roberts et al. (2014b), that the results should not only be dependent upon the data, but also upon human judgment.

Trends of Sub-Topics

Trends of sub-topics reflect the degree of researchers' interest in a research topic. As stated in the research methodology section, the trend of a topic includes two aspects: the change of prevalence over time and the change of research contents over time.

The change of prevalence of a topic can be obtained using STM. As the section result stated, STM functions in a manner similar to standard regression analysis, where topic-proportions are the outcome variable and publication time is the covariate (Roberts et al., 2014a, 2018). The result indicated that only the estimate of topic 1 had statistical significance ($p<.05$) with a positive value (see Table 2, Figure 2), which meant that only the prevalence of topic 1 had been showing an upward trend. Therefore, it can be concluded that only topic 1 might attract increasing interest from researchers among all topics mentioned in the previous section.

Through the analysis of co-occurrence at the document level, the correlation within the topics can be found (see Figure 1, Roberts et al., 2014a). That is, the topics, linked together, would frequently be discussed in the same document. Therefore, theoretically speaking, topics linked together may show the same trend for prevalence. To illustrate this, if two topics (topic A and topic B) are linked together, and topic A shows an upward trend while topic B shows a downward trend, topic B might attract more interest from researchers due to the upward trend of topic A. Thus, it can be inferred that if there is a topic linked with topic 1, even if this topic currently shows a downward trend, an upward trend of this topic is regarded as probable, due to the sustained upward trends of

https://doi.org/10.33225/jbse/20.19.551

topic 1. In fact, there is no edge linking other topics with topic 1 in Figure 1. In sum, in this study, the focus of topic 1 was regarded as the trends of research on students' conceptual understanding.

As the above section mentioned, topic 1 focused on the scientific argumentation/reasoning in the construction of students' conceptions. This result can only provide the focus of topic 1 rather than the change of focus over time. Fortunately, STM provides a function to estimate the correlation between the phrases' use within a given topic and publication time using a method that is similar to standard regression analysis (Roberts et al., 2014a, 2018) (see result section, Figure 3).

This study is concerned with trends of research on students' conceptual understanding, and the second period represented these trends. Therefore, only the research contents of the second period are discussed here. As Figure 3 shows, the second period emphasized words including learn (ing, this results from the pre-process using lemmatization) progression, natural selection, evolution (shared by both periods), etc.

As the result section mentioned, the more frequently a word is used, the higher the probability of it belonging to a topic (Roberts et al., 2018). As Figure 3 shows, ***evolution*** (shared by both periods), ***natural selection***, and ***learning progression,*** are the three largest phrases among all words belonging to the second period of topic 1. Therefore, these three words have the largest probability among all words belonging to the second period of topic 1 in Figure 3 (like ***hypothetico*** of topic 1 in the above section), which means that these three phrases all have potential to represent the focus of it. In other aspects, the x coordinate of phrases in the second period remains closer to the right, having more potential to represent the characteristics of the second period. Therefore, compared to evolution and natural selection, learning progression had more potential to explain the trends of research of topic 1. In conclusion, though the words of the second period, except for learning progression, can represent the focus of the second period of topic 1, learning progression can represent the trends of the second period of topic 1 more.

Learning progressions are generally viewed by researchers as conjectural or hypothetical model pathways of learning over periods of time that have been empirically validated (Duschl et al., 2011). As the discussion section stated, topic 1 focused on **scientific argumentation/reasoning.** Taking this result into account, it might be concluded that the research on the development (or pathways) of students' scientific argumentation/reasoning over periods of time may attract more interest from researchers in the foreseeable future.

The Foundations of Research on Students' Conceptual Understanding

The result section listed the 10 most highly cited papers of both periods. As mentioned earlier, there are unique studies in each period and there are identical studies between both periods. In this section, the foundations of the topics were obtained through reviewing these unique or identical papers' changes over time in each period.

The first period included 10 papers, which mainly focused on three aspects of research on students' conceptual understanding: the studies that played an important role in the period when researchers began to focus on the research pertaining to students' conceptual understanding (e.g., term definitions and reviews); the factors (instruction tools or conditions) that could promote students' conceptual understanding; and the investigation of students' conceptions. As for the first aspect, the studies included research about the epistemological and ontological status of misconceptions, preconceptions, alternative conceptions, etc. (Gilbert & Watts, 1983), a review of theoretical and empirical issues in the study of students' conceptual frameworks (Driver & Erickson, 1983), a review of literature-related concept development in adolescent science students (Driver & Easley 1978), and the work of Posner et al. (1982) initiating the research about conceptual change. The second aspect includes the research on the conditions of conceptual change in the classroom (Hewson & Thorley, 1989), the mediators of the process of conceptual change (Pintrich et al., 1993), and the powerful strategies, including concept mapping and Vee diagramming, to help students construct their conceptual understanding (Novak et al., 1984). The third aspect includes research about the investigation of students' conceptions (or children's science, Gilbert et al. 1982, Osborne & Freyberg, 1985; children's ideas, Driver, 1985).

Compared to the first period, the highly cited studies of the second period gave attention to far more than three aspects of research about students' conceptual understanding, namely the nature (or mental process, or mechanism) of the process of students' conceptual understanding (or change, see Disessa, 1993; Duschl & Hamilton, 1992; Vosniadou, 1994), description of the process of students' conceptual understanding (or change, Carey, 1985, Vosniadou & Brewer, 1992), investigation of students' conceptions (Driver, 1994), the mediators of

https://doi.org/10.33225/jbse/20.19.551

ISSN 1648–3898 /Print/
ISSN 2538–7138 /Online/

the process of conceptual change (Pintrich et al., 1993), and the value of the framework of conceptual change for science learning (Duit & Treagust, 2003); the work of Posner et al. (1982) initiated the research about conceptual change. It should be noticed that the document named National Science Education Standards (Council & others, 1996) was not discussed here, because it is not purely designed to develop students' conceptual understanding (though numerous studies referred to it).

Like Berger (1979), Klopfer (1983) stated that science education was still considered a pre-paradigmatic domain around 1980. Therefore, in the research about students' conceptual understanding, which is one of the research topics of science education that no doubt was a pre-paradigmatic domain, this meant that new terms had to be defined. Many of the highly cited studies in the first period (1980–1999) were studies like this (four of ten studies, Driver & Easley, 1978; Driver & Erickson, 1983; Gilbert & Watts, 1983, and Posner et al., 1982). Since the 1980s, science education had experienced impressive development and had become its own domain of research. Most of the highly cited studies in the second period were published in this period, and most of them concerned the description of the process of students' conceptual understanding (or change), or the nature (or mental process, or mechanism) of the process of students' conceptual understanding (five of nine studies, one excluded) (Carey, 1985; Disessa, 1993; Duschl & Hamilton, 1992; Vosniadou, 1994; Vosniadou & Brewer, 1992). That is to say, compared to the studies in the first period, the studies in the second period prefer to focus on the research about the description (or nature, mental process, etc.) of the process of students' conceptual understanding as their research foundations.

The identical aspects between these periods represent the continuity between them. As mentioned above, empirical evidence provided by the research about the investigation of students' conceptions (Driver, 1994; Osborne & Freyberg, 1985) and the research about mediators of the process of conceptual change (Pintrich et al., 1993) were regarded as the research foundations in both periods. In addition, the work of Posner et al. (1982) initiated the research about conceptual change, which was also seen as the foundations in both periods.

Conclusions and Limitations

In this study, with consideration of the information about the semantic cohesion and exclusivity of words to topics, the retrieved articles were divided into 10 sub-topics using STM. Topic 1, **scientific argumentation/reasoning**, was seen to attract more interest from researchers, which was determined using a method similar to standard regression analysis (see discussion section). The foundations of research on students' conceptual understanding in each period were also obtained by reviewing the 10 most highly cited papers of these periods. The conclusions of this study were drawn as follows:

1. There are 10 sub-topics of research on students' conceptual understanding, which include **scientific argumentation/reasoning**, **teachers' pedagogical content knowledge**, **nature of science**, etc.
2. Only topic 1 (scientific argumentation/reasoning, which plays an important role in the construction of students' conceptions) might attract increasing interest from researchers among all topics mentioned above. Specially, the research about the development (or pathways) of students' scientific argumentation/reasoning over periods of time may attract more interest from researchers in the foreseeable future.
3. The result about highly cited studies shows that the studies in the second period (from 2000 until now) tend to regard research about the description (or nature, mental process, etc.) of the process of students' conceptual understanding as the research foundation in comparison to the studies in the first period (before 2000). Further, the empirical evidence provided by the research about the investigation of students' conceptions and the research about mediators of the process of conceptual change has been the foundation of both periods.

There are many advantages to obtaining the trends and foundations of research on students' conceptual understanding using STM. One advantage is that it divides articles into different groups using a method for unsupervised classification rather than a supervised method where the topics are defined by hand-coding a corpus of documents. A second advantage is that it can discover topics and estimate their relationship (in a manner that is similar to standard regression analysis) to document metadata (such as publication date). It is noteworthy that STM has a number of limitations. Besides the complexity of the method and the difficulty in planning survey experiments, the method does not provide a direct quantification of focus or interest when studying the change of research content of topic 1 (see result section, discussion section). Therefore, the analysis

https://doi.org/10.33225/jbse/20.19.551

ISSN 1648–3898 /Print/
ISSN 2538–7138 /Online/

of the change of interest in topic 1 may lead to subjectivity, although this paper reduces subjectivity in further analysis by analyzing the position and size of phrases. Finally, although a recent comparison showed that the identification of topics with STM provided similar results to human coding, the results obtained in this study need to be tested further over time. In conclusion, although there are some limitations to obtaining the trends and foundations using STM, this study highlights the value of using STM to address trends and foundations of research on students' conceptual understanding. Meanwhile, this study holds that the value of STM will be more fully achieved as the studies on students' conceptual understanding become more abundant.

Acknowledgements

This work was supported by the 13th Five-Year Plan on Educational Science Project (YZ2017029) of Shandong Province of China.

References

Aron, A., Aron, E., & Coups, E. (2013). *Statistics for psychology* (6th ed.). Pearson Education.
American Psychological Association (APA). (2009). *Publication Manual of the American Psychological Association* (6th ed.). American Psychological Association.
Bakkalbasi, N., Bauer, K., Glover, J., & Wang, L. (2006). Three options for citation tracking. *Biomedical Digital Libraries, 3*(1), 7. https://doi.org/10.1186/1742-5581-3-7
Banerjee, A. C. (1991). Misconceptions of students and teachers in chemical equilibrium. *International Journal of Science Education, 13*(4), 487–494. https://doi.org/10.1080/0950069910130411
Banks, G. C., Woznyj, H. M., Wesslen, R. S., & Ross, R. L. (2018). A review of best practice recommendations for text analysis in R (and a user-friendly app). *Journal of Business and Psychology, 33*(4), 445–459. https://doi.org/10.1007/s10869-017-9528-3
Basili, P. A., & Sanford, J. P. (1991). Conceptual change strategies and cooperative group work in chemistry. *Journal of Research in Science Teaching, 28*(4), 293–304. https://doi.org/10.1002/tea.3660280403
Benoit, K., & Matsuo, A. (2019). spacyr. Wrapper to the 'spaCy' 'NLP' Library. https://CRAN.R-project.org/package=spacyr
Berger, C. F. (1979). What are the implications of paradigm-based research for science education research? *Journal of Research in Science Teaching, 16*(6), 517–521. https://doi.org/10.1002/tea.3660160605
Berry, L. L., & Parasuraman, A. (1993). Building a new academic field: The case of services marketing. *Journal of Retailing, 69*(1), 13–60. https://doi.org/10.1016/S0022-4359(05)80003-X
Bezzi, A. (1996). Use of repertory grids in facilitating knowledge construction and reconstruction in geology. In *Journal of Research in Science Teaching, 33*(2), 179–204. https://doi.org/10.1002/(SICI)1098-2736(199602)33:2<179::AID-TEA4>3.0.CO;2-R
Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet allocation. *Journal of Machine Learning Research, 3*(Jan), 993–1022.
Bohr, J., & Dunlap, R. E. (2018). Key Topics in environmental sociology, 1990–2014. Results from a computational text analysis. *Environmental Sociology, 4*(2), 181–195. https://doi.org/10.1080/23251042.2017.1393863
Brumby, M. N. (1984). Misconceptions about the concept of natural selection by medical biology students. *Science Education, 68*(4), 493–503. https://doi.org/10.1002/sce.3730680412
Carey, S. (1985). *Conceptual change in childhood*. MIT press.
Chandelier, M., Steuckardt, A., Mathevet, R., Diwersy, S., & Gimenez, O. (2018). Content analysis of newspaper coverage of wolf recolonization in France using structural topic modeling. *Biological Conservation, 220*, 254–261. https://doi.org/10.1016/j.biocon.2018.01.029
Chang, Y-H, Chang, C-Y, & Tseng, Y-H (2010). Trends of science education research. An automatic content analysis. *Journal of Science Education and Technology, 19(*4), 315–331. https://doi.org/10.1007/s10956-009-9202-2
Cho, H-H., Kahle, J. B., & Nordland, F. H. (1985). An investigation of high school biology textbooks as sources of misconceptions and difficulties in genetics and some suggestions for teaching genetics. *Science Education, 69(*5), 707–719. https://doi.org/10.1002/sce.3730690512
Claesgens, J., Scalise, K., Draney, K., Wilson, M., & Stacy, A. (2002). *Perspective of a Chemist: A Framework to Promote Conceptual Understanding of Chemistry*. Annual Meeting of the American Educational Research Association, New Orleans, LA.
Cuccurullo, C., Aria, M., & Sarto, F. (2016). Foundations and trends in performance management. A twenty-five year bibliometric analysis in business and public administration domains. *Scientometrics, 108*(2), 595–611. https://doi.org/10.1007/s11192-016-1948-8
Disessa, A. A. (1993). Toward an epistemology of physics. *Cognition and Instruction, 10*(2-3), 105–225. https://doi.org/10.1080/07370008.1985.9649008
Driver, R. (1985). *Children's ideas in science*. McGraw-Hill Education (UK).
Driver, R., Rushworth, P., Squires, A., & Wood-Robinson, V. (Eds.). (2005). *Making sense of secondary science: Research into children's ideas*. Routledge.
Driver, R., & Easley, J. (1978). Pupils and paradigms: A review of literature related to concept development in adolescent science students. *Studies in Science Education*, *5*, 61-84. https://doi.org/10.1080/03057267808559857

https://doi.org/10.33225/jbse/20.19.551

ISSN 1648–3898 /Print/
ISSN 2538–7138 /Online/

Driver, R., & Erickson, G. (1983). Theories-in-action: Some theoretical and empirical issues in the study of students' conceptual frameworks in science. *Studies in Science Education, 10*(1), 37–60. https://doi.org/10.1080/03057268308559904

Duit, R., & Treagust, D. F. (2003). Conceptual change. A powerful framework for improving science teaching and learning. *International Journal of Science Education, 25*(6), 671–688. https://doi.org/10.1080/09500690305016

Dukes, R. J., & Strauch, K. (1984). Trends in physics teaching. *The Physics Teacher, 22*(7), 446–448. https://doi.org/10.1119/1.2341613

Duschl, R. A., & Hamilton, R. J. (1992). *Philosophy of science, cognitive psychology, and educational theory and practice*. State University of New York Press.

Duschl, R., Maeng, S., & Sezen, A. (2011). Learning progressions and teaching sequences: A review and analysis. *Studies in Science Education*, *47*(2), 123-182. https://doi.org/10.1080/03057267.2011.604476

Dybowski, T. P., & Adämmer, P. (2018). The economic effects of US presidential tax communication. Evidence from a correlated topic model. *European Journal of Political Economy, 55*, 511–525. https://doi.org/10.1016/j.ejpoleco.2018.05.001

Fischer, F., Kollar, I., Ufer, S., Sodian, B., Hussmann, H., Pekrun, R., Neuhaus, B., Dorner, B., Pankofer, S., Fischer, M., & Strijbos, J. W. (2014). Scientific reasoning and argumentation. Advancing an interdisciplinary research agenda in education. *Frontline Learning Research, 2*(3), 28–45. https://doi.org/10.14786/flr.v2i2.96

Gallegos, L., Jerezano, M. E., & Flores, F. (1994). Preconceptions and relations used by children in the construction of food chains. *Journal of Research in Science Teaching, 31*(3), 259–272. https://doi.org/10.1002/tea.3660310306

Geese, L. (2019). Immigration-related speechmaking in a party-constrained parliament: Evidence from the "refugee crisis" of the 18th German Bundestag (2013–2017). *German Politics*, 1–22. https://doi.org/10.1080/09644008.2019.1566458

Gilbert, J. K., Osborne, R. J., & Fensham, P. J. (1982). Children's science and its consequences for teaching. *Science Education, 66(*4), 623–633. https://doi.org/10.1002/SCE.3730660412

Gilbert, J. K., & Watts, D. M. (1983). Concepts, misconceptions and alternative conceptions: Changing perspectives in science education. *Studies in Science Education, 10*(1), 61–98. https://doi.org/10.1080/03057268308559905

Gilbert, J. K., Bulte, A. M. W., & Pilot, A. (2011). Concept development and transfer in context-based science education. *International Journal of Science Education, 33*(6), 817–837. https://doi.org/10.1080/09500693.2010.493185

Grimmer, J., & Stewart, B. M. (2013). Text as data: The promise and pitfalls of automatic content analysis methods for political texts. *Political Analysis, 21*(3), 267–297. https://doi.org/10.1093/pan/mps028

Ha, M., Haury, D. L., & Nehm, R. H. (2012). Feeling of certainty: Uncovering a missing link between knowledge and acceptance of evolution. *Journal of Research in Science Teaching, 49*(1), 95–121. https://doi.org/10.1002/tea.20449

Hewson, M. G., & Hewson, P. W. (1983). Effect of instruction using students' prior knowledge and conceptual change strategies on science learning. *Journal of Research in Science Teaching, 20*(8), 731–743. https://doi.org/10.1002/tea.3660200804

Hewson, P. W., & Thorley, N. R. (1989). The conditions of conceptual change in the classroom. *International Journal of Science Education, 11*(5), 541–553. https://doi.org/10.1080/0950069890110506

de Jong, O. (2007). Trends in western science curricula and science education research: A bird's eye view. *Journal of Baltic Science Education, 6*(1), 15–22.

Kang, S., Scharmann, L. C., & Noh, T. (2004). Reexamining the role of cognitive conflict in science concept learning. *Research in Science Education, 34*(1), 71–96. https://doi.org/10.1023/B:RISE.0000021001.77568.B3

Klopfer, L. E. (1983). Editorial: Research and the crisis in science education. *Science Education, 67*(3), 283–284. https://doi.org/10.1002/sce.3730670302

Lavoie, D. R. (1999). Effects of emphasizing hypothetico-predictive reasoning within the science learning cycle on high school students' process skills and conceptual understandings in biology. *Journal of Research in Science Teaching, 36*(10), 1127–1147. https://doi.org/10.1002/(SICI)1098-2736(199912)36:10<1127::AID-TEA5>3.0.CO;2-4

Lawson, A. E. (1985). A review of research on formal reasoning and science teaching. *Journal of Research in Science Teaching, 22*(7), 569–617. https://doi.org/10.1002/tea.3660220702

Lawson, A. E., & Worsnop, W. A. (1992). Learning about evolution and rejecting a belief in special creation. Effects of reflective reasoning skill, prior knowledge, prior belief and religious commitment. *Journal of Research in Science Teaching, 29*(2), 143–166. https://doi.org/10.1002/TEA.3660290205

Lawson, A. (2003). The nature and development of hypothetico-predictive argumentation with implications for science teaching. *International Journal of Science Education, 25*(11), 1387–1408. https://doi.org/10.1080/0950069032000052117

Lee, M.-H., Wu, Y.-T., & Tsai, C-C (2009). Research trends in science education from 2003 to 2007. A content analysis of publications in selected journals. *International Journal of Science Education, 31*(15), 1999–2020. https://doi.org/10.1080/09500690802314876

Lin, J.-Y. (2007). Responses to anomalous data obtained from repeatable experiments in the laboratory. *Journal of Research in Science Teaching, 44*(3), 506–528. https://doi.org/10.1002/TEA.20125

Lin, T.-C., Lin, T.-J., & Tsai, C.-C. (2014). Research trends in science education from 2008 to 2012. A systematic content analysis of publications in selected journals. *International Journal of Science Education, 36*(8), 1346–1372. https://doi.org/10.1080/09500693.2013.864428

Lin, T., Lin, T., Potvin, P., & Tsai, C. (2019). Research trends in science education from 2013 to 2017: A systematic content analysis of publications in selected journals. *International Journal of Science Education, 41*(3), 367–387. https://doi.org/10.1080/09500693.2018.1550274

National Research Council. (1996). *National science education standards*. National Academies Press.

Niaz, M. (1988). Manipulation of M demand of chemistry problems and its effect on student performance: A neo-piagetian study. *Journal of Research in Science Teaching, 25*(8), 643–657. https://doi.org/10.1002/tea.3660250804

https://doi.org/10.33225/jbse/20.19.551

ISSN 1648–3898 /Print/
ISSN 2538–7138 /Online/

Niaz, M. (1995). Cognitive conflict as a teaching strategy in solving chemistry problems: A dialectic–constructivist perspective. *Journal of Research in Science Teaching, 32*(9), 959–970. https://doi.org/10.1002/tea.3660320907

Novak, J. D., & Gowin, D. B.. (1984). *Learning how to learn*. Cambridge University Press.

Olympiou, G., & Zacharia, Z. C. (2012). Blending physical and virtual manipulatives: An effort to improve students' conceptual understanding through science laboratory experimentation. *Science Education, 96*(1), 21–47. https://doi.org/10.1002/sce.20463

Osborne, J. (2010). Arguing to learn in science: The role of collaborative, critical discourse. *Science (New York, N.Y.), 328*(5977), 463–466. https://doi.org/10.1126/science.1183944

Osborne, R., & Freyberg, P. (1985). Learning in science. The implications of children's science: ERIC.

Park, J. (2006). Modelling Analysis of Students' Processes of Generating Scientific Explanatory Hypotheses. *International Journal of Science Education, 28*(5), 469–489. https://doi.org/10.1080/09500690500404540

Peel, A., Sadler, T. D., & Friedrichsen, P. (2019a). Learning natural selection through computational thinking: Unplugged design of algorithmic explanations. *Journal of Research in Science Teaching, 56*(7), 983–1007. https://doi.org/10.1002/TEA.21545

Peel, A., Zangori, L., Friedrichsen, P., Hayes, E., Sadler, T. (2019b). Students' model-based explanations about natural selection and antibiotic resistance through socio-scientific issues-based learning. *International Journal of Science Education, 41*(4), 510–532. https://doi.org/10.1080/09500693.2018.1564084

Pines, A. L., & West, L. H. T. (1986). Conceptual understanding and science learning: An interpretation of research within a sources-of-knowledge framework. *Science Education, 70*(5), 583–604. https://doi.org/10.1002/sce.3730700510

Pintrich, P. R., Marx, R. W., & Boyle, R. A. (1993). Beyond cold conceptual change: The role of motivational beliefs and classroom contextual factors in the process of conceptual change. *Review of Educational Research, 63*(2), 167–199. https://doi.org/10.3102/00346543063002167

Posner, G. J., Strike, K. A., Hewson, P. W., Gertzog, W. A. (1982). Accommodation of a scientific conception: Toward a theory of conceptual change. *Science Education, 66*(2), 211–227. https://doi.org/10.1002/SCE.3730660207

Reich, J., Tingley, D., Leder-Luis, J., Roberts, M. E., & Stewart, B. (2015). Computer-assisted reading and discovery for student generated text in massive open online courses. *Journal of Learning Analytics, 2*(1), 156–184. https://doi.org/10.18608/jla.2015.21.8

Reif, F. (1995). Millikan Lecture 1994. Understanding and teaching important scientific thought processes. *American Journal of Physics, 63*(1), 17–32. https://doi.org/10.1119/1.17764

Roberts, M. E., Stewart, B. M., & Tingley, D., (2014a). STM: R package for structural topic models. *Journal of Statistical Software, 10*(2), 1–40. https://doi.org/10.18637/jss.v091.i02

Roberts, M. E., Stewart, B. M., Tingley, D., Lucas, C., Leder-Luis, J., Gadarian, S. K., Albertson, B. & Rand, D. G. (2014b). Structural topic models for open-ended survey responses. *American Journal of Political Science, 58*(4), 1064–1082. https://doi.org/10.1111/ajps.12103

Roberts, M. E., Stewart, B. M., & Tingley, D. (2018). *STM: R package for structural topic models*. http://www.structuraltopicmodel.com

Romine, W. L., Walter, E. M., Bosse, E., & Todd, A. N. (2017). Understanding patterns of evolution acceptance: A new implementation of the Measure of Acceptance of the Theory of Evolution (MATE) with Midwestern university students. *Journal of Research in Science Teaching, 54*(5), 642–671. https://doi.org/10.1002/TEA.21380

Salinas, I. (2009). Learning progressions in science education: Two approaches for development. In *Learning Progressions in Science (LeaPS) Conference*, Iowa City, IA.

Shirokanova, A., & Silyutina, O. (2018) Internet regulation: A text-based approach to media coverage. In: Alexandrov, D., Boukhanovsky, A., Chugunov, A., Kabanov, Y., & Koltsova, O. (Eds), *Digital transformation and global society. DTGS 2018. Communications in computer and information science* (Vol. 858). Springer. https://doi.org/10.1007/978-3-030-02843-5_15

Simpson, W. D., & Marek, E. A. (1988). Understandings and misconceptions of biology concepts held by students attending small high schools and students attending large high schools. *Journal of Research in Science Teaching, 25*(5), 361–374. https://doi.org/10.1002/tea.3660250504

Stavy, R. (1991). Using analogy to overcome misconceptions about conservation of matter. *Journal of Research in Science Teaching, 28*(4), 305–313. https://doi.org/10.1002/tea.3660280404

Taber, K. S. (2001). Shifting sands: A case study of conceptual development as competition between alternative conceptions. *International Journal of Science Education, 23*(7), 731–753. https://doi.org/10.1080/09500690010006572

Taber, K. S. (2014). *Student thinking and learning in science: Perspectives on the nature and development of learners' ideas*. Routledge.

Taber, K. S. (2015a). Alternative conceptions/frameworks/misconceptions. In: Gunstone, R. (Ed.), *Encyclopedia of science education*. Springer. https://doi.org/10.1007/978-94-007-2150-0_88

Taber, K. S. (2015b). Prior knowledge. In: Gunstone, R. (Ed.), *Encyclopedia of science education*. Springer. https://doi.org/10.1007/978-94-007-2150-0_483

Tsai, C.-C., & Chang, C.-Y. (2005). Lasting effects of instruction guided by the conflict map. Experimental study of learning about the causes of the seasons. *Journal of Research in Science Teaching: The Official Journal of the National Association for Research in Science Teaching, 42*(10), 1089–1111. https://doi.org/10.1002/tea.20089

Tsai, C.-C., & Lydia Wen, M. (2005). Research and trends in science education from 1998 to 2002. A content analysis of publication in selected journals. *International Journal of Science Education, 27*(1), 3–14. https://doi.org/10.1080/0950069042000243727

Vosniadou, S. (1994). Capturing and modeling the process of conceptual change. *Learning and Instruction, 4*(1), 45–69. https://doi.org/10.1016/0959-4752(94)90018-3

https://doi.org/10.33225/jbse/20.19.551

ISSN 1648–3898 /Print/
ISSN 2538–7138 /Online/

Vosniadou, S., & Brewer, W. F. (1992). Mental models of the earth: A study of conceptual change in childhood. *Cognitive Psychology, 24*(4), 535–585. https://doi.org/10.1016/0010-0285(92)90018-W
White, R. (1997). Trends in research in science education. *Research in Science Education,* 27(2), 215–221. https://doi.org/10.1007/BF02461317
Zangori, L., Peel, A., Kinslow, A., Friedrichsen, P., & Sadler, T. D. (2017). Student development of model-based reasoning about carbon cycling and climate change in a socio-scientific issues unit. *Journal of Research in Science Teaching, 54*(10), 1249–1273. https://doi.org/10.1002/TEA.21404
Zhang, W., Zhang, W., Yu, B., & Zhao, L. (2015). Knowledge map of creativity research based on keywords network and co-word analysis, 1992–2011. *Quality & Quantity, 49*(3), 1023–1038. https://doi.org/10.1007/s11135-014-0032-9

Received: *April 04, 2020* Accepted: *July 29, 2020*

Cite as: Mi, S., Lu, S., & Bi, H. (2020). Trends and foundations in research on students' conceptual understanding in science education: A method based on the structural topic model. *Journal of Baltic Science Education*, *19*(4), 551-568. https://doi.org/10.33225/jbse/20.19.551

Shuaishuai Mi	PhD Student, College of Chemistry, Chemical Engineering and Materials Science, Shandong Normal University, Ji'nan, Shandong 250014, China. E-mail: mishuaishuai@stu.sdnu.edu.cn ORCID: https://orcid.org/0000-0001-5308-1279
Shanshan Lu	Doctor of Education, Associate Professor, College of Chemistry, Chemical Engineering and Materials Science, Shandong Normal University, Ji'nan, Shandong 250014, China. E-mail: 615046@sdnu.edu.cn ORCID: https://orcid.org/0000-0002-5492-6285
Hualin Bi *(Corresponding author)*	Doctor of Education, Professor, College of Chemistry, Chemical Engineering and Materials Science, Shandong Normal University, Ji'nan, Shandong 250014, China. E-mail: bihl@qfnu.edu.cn ORCID: http://orcid.org/0000-0002-2244-4839