Science Education International
37(1)_2026,17-27
https://doi.org/10.33828/sei.v37.i1.2

International Council of Association for Science Education

ORIGINAL ARTICLE

Making the Invisible Visible: A Systematic Review of Mental Models in Chemistry Education

Rüdiger Tiemann*, Anna-Lena Nitsche

Department of Chemistry, Humboldt-Universität zu Berlin, Chemistry Education Research, Berlin, Germany

***Corresponding Author:** ruediger.tiemann@hu-berlin.de

ABSTRACT

This systematic literature review analyzes 22 studies on mental models in chemistry education published between 2013 and 2022. Findings reveal a growing research interest, particularly since 2015, with a strong geographical concentration in Asia – most notably Indonesia – likely linked to the implementation of the 2013 curriculum (K-13) (Machali, 2014). The reviewed studies vary methodologically: More than half were basic investigations exploring the structure, composition, and misconceptions of mental models (Seel, 2014), while others implemented interventions using dynamic simulations (Akaygun, 2016a; Akaygun, 2016b) or evaluated diagnostic tools (Chiang and Chiu, 2015). Qualitative and mixed-method approaches dominate, especially semi-structured interviews (Döring and Bortz, 2016; Greca and Moreira, 1997). The studies primarily address fundamental chemistry concepts such as bonding, atomic structure, and acid-base reactions (Yildirir and Demirkol, 2018). Results indicate that dynamic visualizations, attention to the three representational levels (Albaiti et al., 2022; Murni et al., 2022), and motivational factors positively influence mental model development, while misconceptions impede it (Putri and Wiyarsi, 2022; Buckley, 2000). The review concludes that methodological diversity is essential for a comprehensive understanding. Future research should refine theoretical frameworks, evaluate innovative digital approaches, and further investigate mechanisms influencing mental model formation to enhance chemistry learning outcomes.

KEY WORDS: Chemistry education, mental models, science, technology, engineering, and mathematic education, systematic literature review

INTRODUCTION

When students wonder why sugar disappears in tea or why iron rusts, they often see the result but miss the hidden processes. Models help fill this gap by making the invisible visible, the abstract tangible, and the complex understandable. From the geocentric worldview to the double helix of DNA, models have always served as essential thinking tools to simplify reality and make it accessible (Seel, 2017). Today, they are central to research and indispensable for teaching and learning (Harrison and Treagust, 2000), with model competence considered a key goal in science, technology, engineering, and mathematic education (Seel, 2014; Seel, 2017).

In classrooms, students should actively construct and apply models to understand phenomena (Cobb et al., 2000; Clement, 2000). Conceptual models organize knowledge (Norman, 1983), while mental models – internal representations – enable reasoning and predictions (Johnson-Laird, 2004; Nitz and Fechner, 2018). Although mental models are not directly observable (Seel, 2017), understanding them is vital for effective learning.

Chemistry highlights this challenge. Its phenomena span three levels: The visible macroscopic substances, the microscopic particles, and the symbolic formulas (Johnstone, 1991). Linking these levels demands model thinking, and mental models serve as bridges that make concepts meaningful and actionable (Johnstone, 1991; Krumm and Scharf, 1992).

This review examines how mental models in chemistry education have been studied over the past decade. Across a systematic database search, studies were identified, coded, and analyzed. Findings are presented by thematic focus and methodology, trends are discussed, and directions for future research are outlined.

MENTAL MODELS

A Coherent Overview

Mental models are internal representations that individuals use to interpret complex phenomena, make predictions, and plan actions (Greca and Moreira, 2000; Seel, 1999). They differ from conceptual models, which serve as external representations such as diagrams or formulas (Norman, 1983; Seel, 2014). Ideally, both complement each other, yet discrepancies often impede learning. Although flexible and incomplete, mental models can stabilize through repeated use (Halford et al., 1998; Seel, 2014).

Research on mental models originates with Ritchie and Craik (1943), who emphasized analogy formation and prediction (Cárdenas-Figueroa and Navairo, 2020). Johnson-Laird

(1983) defined mental models as dynamic representations mediating between images and propositions (Greca and Moreira, 2000). Despite extensive interdisciplinary research, consistent definitions remain lacking (Nitz and Fechner, 2018; Thagard, 2010).

In science education, mental models form the foundation of conceptual understanding. Learners construct and refine models through visualizations, experiments, and targeted instruction (Buckley, 2000; Clement, 2008; Seel, 2017). Prior knowledge strongly influences model development (Greca and Moreira, 2000; Larson et al., 2012), yet rote learning often leads to fragmented understanding (Seel, 2014). Effective teaching thus requires active model construction, reflection, and the correction of misconceptions (Gilbert and Boulter, 1998; Ifenthaler and Seel, 2011).

Methodologically, investigating mental models is challenging due to their internal and dynamic nature (Gobert and Buckley, 2000). Researchers employ drawings, concept maps, and verbal protocols to infer model structures and knowledge application (Glynn, 1997; Rumelhart et al., 1986; Seel, 2017). Studies range from structural analyses to interventions using simulations and digital learning tools (Gobert and Buckley, 2000; Seel, 2014; Seel, 2017).

Overall, mental models represent a crucial interface between cognition, learning, and instruction. Effective science teaching depends on recognizing and developing students' models, while research requires integrating qualitative and quantitative approaches to capture their complexity (Greca and Moreira, 2000; Seel, 2014).

Assessment Methods

For years, researchers have debated how to identify mental models and represent their development. Some methods visualize models to capture their present state or difficulties, while others use technology to record changes over time (Seel, 2014). Two main approaches dominate this field: graphical techniques that map structural relations and linguistic analyses based on verbal data, such as think-aloud protocols (ibid.).

Because mental models are theoretical constructs (Seel, 2017), all methods rest on assumptions derived from behavior, verbalizations, or computer simulations (Moreira, 1996). Ideally, diagnosis occurs within authentic problem-solving contexts and longitudinal designs with valid quantitative data (Seel, 1999). Yet, mental models are often incomplete and not consciously accessible (Moreira, 1996), and defining suitable dependent variables remains challenging (Seel, 1999).

Earlier research emphasized open verbalization (Seel, 2014; Greca and Moreira, 1997), though its validity was later questioned (Halford, 1993). This prompted the development of additional techniques such as tests, questionnaires, rating scales, time measures, drawings, and eye-tracking (Seel, 1999). These can be grouped into experimental, protocol-based, and computer-assisted approaches (ibid.). Experimental methods, following Mayer (1989), infer model quality from task performance after training and varied learning phases (Seel, 1999). Protocol analyses rely on verbal data but offer limited validity (Seel, 1999; Norman, 1983).

Since the 2000s, computer-assisted approaches have gained prominence, supporting the organization, visualization, and simulation of knowledge structures (Jonassen, 1999; Seel, 2017). They make complex thought processes more transparent and accessible for research and instruction.

Categorization of Mental Models

Due to their incomplete nature and lack of clear boundaries, the representation of mental models is as problematic as their measurement. In research, they are either created inductively from consistent patterns in the data or derived deductively from theories. Despite differing terminology, cognitive and mental models are usually divided into four types (Kurdziel et al., 2003):

- Naive mental models: Intuitive and unconscious ideas.
- Unstable models: Inaccurate, incomplete, but already in flux.
- Conceptual frameworks: Organized, stable, and frequently used models, typical for beginners.
- Conceptual model: Requires expert knowledge and is considered scientifically correct.

According to Clement (2000), students' mental models can also be divided into an initial alternative model, one to three intermediate models, and a target model.

Educational research hopes to be able to demonstrate a direct relationship between conceptual and mental models as a result of model-based teaching or targeted interventions (Kurdziel et al., 2003).

Research Question

The literature on mental models shows that physics, particularly astronomy (e.g., Vosniadou and Brewer 1992; Vosniadou and Brewer, 1994), is the most frequently studied discipline in teaching research. This is confirmed by Benzer and Ünal (2021) in their review on models and modeling in science education in Turkey. This paper aims to address this research gap by asking:

What insights and trends emerge from English-language research on chemistry education published between 2013 and 2022 regarding students' mental models?

METHOD

With increasing digitalization, publications are becoming more accessible online, making systematic literature reviews increasingly important in research (Harris et al., 2014; Moher et al., 2009). According to the cochrane collaboration, a systematic review summarizes primary literature on a specific research question. Relevant evidence is identified, selected, summarized, and evaluated according to predefined criteria, ensuring transparency and minimizing potential biases (Harris et al., 2014; Pollock and Berge, 2018).

Literature Research Process

According to Xiao and Watson (2017), a systematic literature review proceeds in three phases: Planning, implementation, and reporting. In the planning phase, the relevance of the topic is assessed, the research question formulated, and inclusion and exclusion criteria defined to ensure validity (Harris et al., 2014; Xiao and Watson, 2017).

During implementation, targeted database searches are conducted, supplemented by backward and forward searches, followed by screening and selection of relevant studies (Kitchenham and Charters, 2007). The reporting phase involves categorizing, analyzing, synthesizing, and transparently presenting the results (Xiao and Watson, 2017).

Xiao and Watson (2017) summarize this process in eight steps: Problem formulation; development and validation of the protocol to reduce subjectivity (Gates, 2002; Gomersall et al., 2015); literature search and definition of search terms; study selection using screening and PRISMA (Harris et al., 2014); quality assessment; data extraction (inductive or deductive); analysis and synthesis; and final reporting (Okoli and Schabram, 2010).

By following this structured approach, systematic reviews maintain transparency, traceability, and comparability, forming a reliable basis for scientific conclusions.

Limitations and Sources of Error in a Literature Review

Although various approaches to conducting a literature review exist in the literature to make the process transparent and consistent, Moher et al. (2009) identified numerous quality deficiencies in literature reviews in a comprehensive review.

A common problem is that the research question is too broad (Xiao and Watson, 2017). This leads to a large amount of data, which makes the research confusing (ibid.). Subsequent narrowing requires a readjustment of the inclusion and exclusion criteria and, as a rule, a new literature search (ibid.). There is also a risk that relevant data sets will not be taken into account (ibid.). In a cross-national literature search, cultural and linguistic differences must be adapted to the search terminology (Pollock and Berger, 2018). For example, "mental models" are used synonymously with "external representation" in some articles, while in another context, such as "(students) performance," are addressed.

METHOD APPLICATION

Literature Screening

Several databases were used to create a comprehensive database and improve the literature review. The search was conducted in English and focused exclusively on electronic databases. Both general and subject-specific databases were included: SCOPUS, ERIC, APA Psycinfo, and PRIMUS. ERIC specializes in educational research, while Psycinfo offers literature on psychology. SCORPUS and PRIMUS are more general databases that cover a broader spectrum and thus also include interdisciplinary studies or studies that cannot be clearly assigned to educational science or psychology. This selection ensured that a broad spectrum of relevant publications was taken into account to create a sound basis for further analysis and conclusions.

Search terms and keywords

For the initial unsystematic literature review, initial keywords were collected by conducting a forward and backward search of the literature. The search terms were divided into the subject of the study and the area of investigation to specify what is being investigated and in which subject area the investigations are taking place. The following Table 1 shows the first keywords collected in the unsystematic literature search in descending order of relevance from top to bottom.

Table 1: Unsystematically collected search terms and keywords

Subject of investigation	Area of investigation
Mental model/mental models	Science education
Mental representation	Chemical (education)
External representation	Chemistry
Model-based learning	Chemical education research
Thought models	

This search in four databases yielded 950 hits, which had to be further narrowed down. The terms "science" and "science education" were too vague. "Chemistry" or "chemical" were sufficient to find literature on chemical education. The search terms were limited to mental model, mental representation, and external representation, as others did not yield clear results. The search was expanded to include the term "students" to further narrow down the literature.

The search was conducted using the search terms and keywords specified in Table 2, sorted by relevance from top to bottom. Boolean operators, such as AND and OR, were used to link and combine the terms.

Table 2: Restrictive search terms and keywords

Subject of investigation	Area of investigation
Mental model:mental models	Chemical (education)
Mental representation	Chemistry
External representation	

The following Table 3 shows the databases used, the search terms, and the respective hits. The term "education" was added for the search in the PRIMUS database.

Inclusion and exclusion criteria for the literature

When selecting the primary studies, specific inclusion and exclusion criteria were established to identify literature relevant to the research question in the 554 data sets.

Inclusion criteria:

1. Studies that have a clear connection to chemistry education research and aim to make new contributions to research in the field of chemistry education.

2. Studies that specifically examine and/or explicitly consider (forms of) mental models.
3. Studies published between 2013 and 2022.
4. Studies that deal with research in schools or in the classroom.
5. Studies that exclusively involve students in grades five to a maximum of 12 or 13 (depending on the state‘s graduation level).
6. Studies that were written and published in English.

Exclusion criteria:
1. Studies that deal exclusively with theory without conducting empirical research.
2. Studies that review other research rather than generating new findings.

This ensured that only relevant primary studies were selected that deal with the investigation of mental models in the context of chemistry education research and can provide new insights and contributions in this area.

Study Selection through Screening

To ensure that the research process is valid and reliable, individual steps and decisions are documented and recorded in the following PRISMA flowchart (Figure 1). The phases of study selection are explained in more detail below.

Identification

After identifying 554 studies and defining the search criteria, duplicates were removed. Despite using EndNote 21 (Clarivate, 2023) for duplicate detection, some duplicates remained, which are sorted out manually. Later, occasionally discovered additional duplicates, for example, due to changes in the order of authors, were also identified. In total, 178 duplicates were removed.

Table 3: Results of the systematic literature search using databases, time period, search terms, and hits

Database	Period	Free/full-text search	Hits
PRIMUS	2013–2022	(mental model) AND (chemistry or chemical) AND (students) AND “education”	257
		(mental representation) AND (chemistry or chemical) AND (students) AND “education”	
		(external representation) AND (chemistry or chemical) AND (students) AND “education”	
SCOPUS	2013–2022	(mental model) AND (chemistry or chemical) AND (students)	180
		(mental representation) AND (chemistry or chemical) AND (students)	
		(external representation) AND (chemistry or chemical) AND (students)	
ERIC	Past 10 years	(mental model) AND (chemistry or chemical) AND (students)	90
		(mental representation) AND (chemistry or chemical) AND (students)	
		(external representation) AND (chemistry or chemical) AND (students)	
Psycinfo	2013–2022	(mental model) AND (chemistry or chemical) AND (students)	27
		(mental representation) AND (chemistry or chemical) AND (students)	
		(external representation) AND (chemistry or chemical) AND (students)	
Total			554

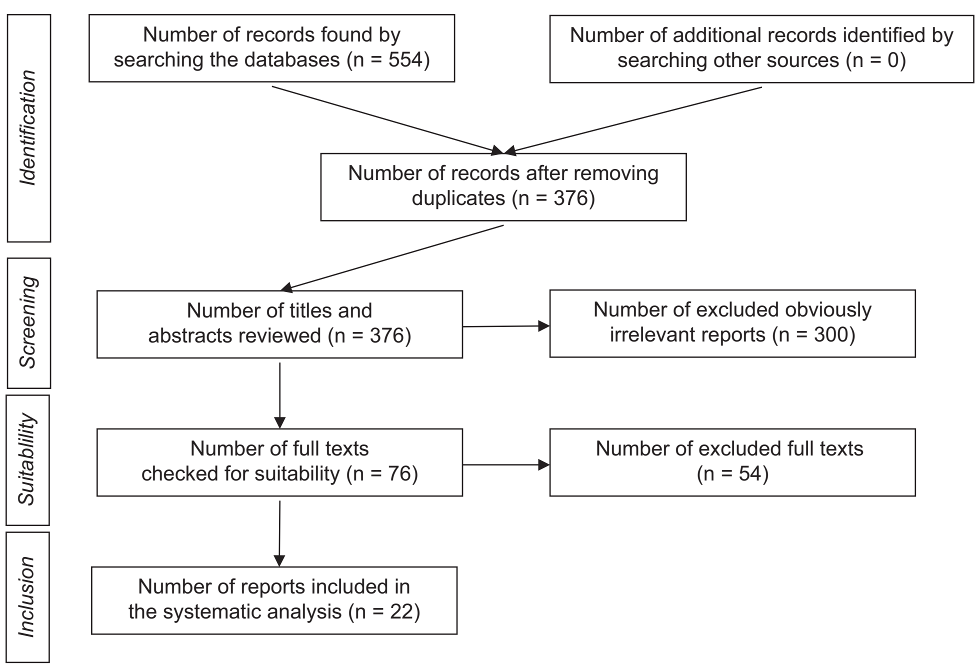

Figure 1: PRISMA flow diagram of the present systematic literature review (modified according to: Moher et al., 2009; Pollock and Berge, 2018)

Screening

During the screening phase, 376 studies were reviewed based on their titles and abstracts to assess alignment with the inclusion criteria. Studies with unclear relevance were retained for further examination. Excluded papers were categorized by research area, research objective, publication year, and participant group.

Seventy-three studies were removed because they did not address chemistry or chemistry education (e.g., Almadhoun and Parham-Mocello, 2021; Goh, 2013). Another 109 were excluded as their objectives did not focus on mental models in chemistry education. A further 116 studies involved participants other than students in grades 5–13, such as teachers or university students (e.g., Tümay, 2014; Derkach, 2021). Two studies were excluded for falling outside the defined time frame. In total, 300 studies were excluded during the initial screening phase.

Suitability

After screening, 76 studies were initially reviewed. Two studies were excluded due to a lack of public access, and seven non-English studies were omitted. The remaining papers were examined in detail according to inclusion and exclusion criteria. Three additional exclusion criteria were introduced: research objective, study subjects, and summaries of other studies. Consequently, 24 studies were excluded based on research objectives, 19 due to unclear subjects, and two (Chen et al., 2015; Tasker, 2016) for lacking original research. In total, 54 studies were excluded after full-text analysis, resulting in 22 studies included in this systematic literature review.

Encoding and data extraction

To address the research question, data from 22 selected studies were systematically extracted and coded to ensure comparability. Eleven descriptive criteria guided the analysis, including publication year, topic, methodology, data collection, study objective, and the application of mental model theory.

For each study, relevant information was organized in a summary table and subsequently coded using a combined inductive–deductive approach. Inductive categories emerged from recurring patterns in the data, while deductive coding followed the framework proposed by Wardah and Wiyarsi (2020). The detailed coding procedure is outlined in the results section.

RESULTS

Chronological Distribution of Selected Studies

Figure 2 shows the number publication per year in the period from 2013 to 2022 to illustrate the temporal distribution of the 22 selected studies. It is clear that no relevant studies on the topic could be found in 2013 and 2014. From 2015 onward, at least one study could be found for each year, with a total of five studies published in 2022. The temporal distribution of the studies in the period under review illustrates a trend toward an increase in research activity from 2015 onward, with the exception of 2016, 2019, and 2021, and a peak in 2022.

Geographical Distribution

As can be seen in Table 4, the vast majority (95%) of the studies were conducted in Asia, with only one study (see Zarkadis et al., 2017) originating from Europe. This shows that Europe has made a smaller quantitative contribution to the available research results than Asia. Eleven studies on mental models in chemistry education research originate from various provinces in Indonesia. Three studies originate from Israel and three from Turkey, indicating a certain relevance of these countries to the topic under investigation. The remaining studies are distributed across Taiwan, Malaysia, and Thailand.

Type of source of the selected studies

The type of publication and the title of the source are shown in Table 5. Most of the articles come from the journal Chemistry Education Research and Practice, with a total of three articles. Many international journals were used, such as the International Journal of Learning, Teaching and Educational Research and the Journal of Research in Science Teaching, as well as some national journals such as the Journal of Turkish Science Education and the Indonesian JKT (Jurnal Tadris Kimiya). Some research results were collected in Indonesia and published in Turkish journals, such as the study by Albaiti et al. (2022).

Chemical Topics and Concepts of the Selected Studies

Table 6 shows various chemical concepts and topics related to mental models. The coding was done inductively, taking into account Wardah and Wiyarsi (2020). A total of ten topics and concepts were identified, most of which relate to the physical,

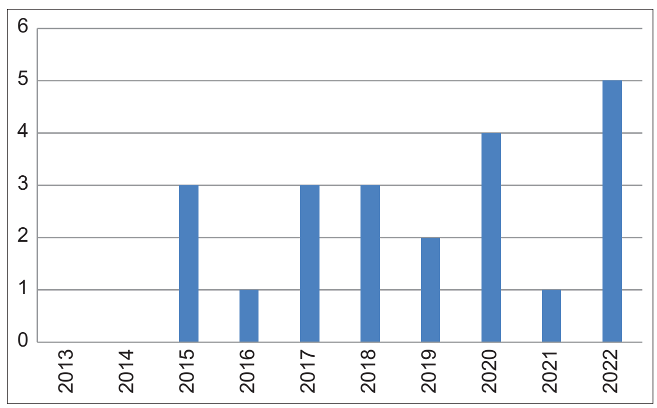

Figure 2: Number of studies per year of publication

Table 4: Geographical distribution worldwide

Country	Number of studies
Indonesien	11
Israel	3
Türkei	3
Taiwan	2
Thailand	1
Malaysia	1
Griechenland	1

inorganic, and analytical areas of chemistry. Due to a lack of clear classifications, the topics were not specified further.

Four studies dealt with atomic structure, such as the construction of atomic orbitals (Sunyono and Sudjarwo, 2018). Another four studies examined mental models of acid-base theories, including those of Arrhenius and Lewis (Putri and Wyiarsi, 2022). Chemical bonds and solubility – e.g., calcium carbonate or sodium chloride in water (Albaiti et al., 2022) – were also frequently investigated. Mole concepts, reaction equations, and physical and chemical changes were each addressed only once.

There is a clear tendency to examine the submicroscopic level according to Johnstone (1991), especially in the case of atomic structure and chemical bonds.

Research Design used in the Selected Studies

The selected studies were evaluated deductively according to Döring and Bortz (2016) and presented in Table 7. This resulted in three types of studies: Basic research, intervention studies, and evaluation studies.

Table 5: Publication medium

Journal	Number of studies
Chemistry Education Research and Practice	3
AIP Conference Proceedings	2
Journal of Research in Science Teaching	2
Journal of Turkish Science Education	2
JTK (Jurnal Tadris Kimiya)	2
Turkish Online Journal of Educational Technology	1
International Journal on Advanced Science, Engineering and Information Technology	1
International Journal of Learning, Teaching and Educational Research	1
Journal of Physics Conference Series	1
Journal of Engineering Science and Technology	1
Asia-Pacific forum on science learning and teaching	1
Computers in Human Behavior	1
Journal of Baltic Science Education	1
Man in India	1
Jurnal Pendidikan IPA Indonesia	1

Table 6: Chemical topics and concepts

Topic/concept	Number of studies
Atomic structure	4
Acids and bases	4
Solubility	3
Chemical bond	2
State of matter	2
Chemical equilibrium	2
Redox-reaction	2
Mole concept	1
Physical and chemical changes	1
Reaction rate	1

Basic research serves to test and gain insights and theories without focusing on intervention (Döring and Bortz, 2016). Intervention research, on the other hand, examines the effect of targeted measures on variables (ibid.), while evaluation research assesses the effectiveness of programs or procedures (ibid.).

More than half of the studies were basic studies that identified mental models without interventions and examined their structure; many of these were conducted in Indonesia. Six intervention studies used specific measures such as simulations (Akaygun, 2016a; Akaygun, 2016b) to record changes in mental models. Four studies can be classified as evaluation studies, for example, to review diagnostic procedures (Chiang and Chiu, 2015). All studies used qualitative methods or mixed methods, that is, a combination of qualitative and quantitative methods.

Data Collection from Selected Studies

A total of 48 methods were used in 22 studies, often in combination (Table 8). Semi-structured interviews were used in 13 studies to investigate mental models. Nine studies used diagnostic tests to identify and evaluate these models. Open-ended questionnaires were also common.

Table 7: Research design

Design	Number of studies
Fundamental research	
Qualitative	10
Mixed-method	2
Interventional study	
Qualitative	2
Mixed-method	4
Evaluation study	
Qualitative	4

Table 8: Methods of data collection

Method	Number of studies
Interview	
Semi-structured	13
Structured	2
Diagnostic assessment	
Mentale models	9
Learning achievement	1
Questionnaire	
Open	7
Observation	
Observation log	3
Video	2
Software	3
Drawing	3
Journaling	2
Minutes	1
Survey	1
Multiple-choice questions	1

Observations, protocols, and digital recordings documented student behavior. Drawings as a standalone method were used in only three studies (see Nasir et al., 2017). Other methods, such as protocols, surveys, or multiple-choice questions, were rarely used. Thirteen methods were identified for gathering information about mental models. The choice of method often depended on the study objectives, which are explained in the next section.

Objectives of the Selected Studies

The results show five main areas, which are summarized in Table 9. The coding was done inductively with reference to Wardah and Wiyarsi (2020).

Ten studies aimed to identify and analyze mental models, mostly using simple designs without pre-post surveys or control groups. Among other things, salt hydrolysis (Kiswandari and Ridwan, 2020), solubility (Rahmi et al., 2017), and atomic structure (Sunyono and Sudjarwo, 2018) were investigated.

Another part of the work examined the effectiveness of interventions, such as static and dynamic simulations (Akaygun, 2016a; Akaygun, 2016b) or the cognitive learning model (Amalia et al., 2018). Pre-post designs or control groups were mostly used here.

Other studies focused generally on misconceptions and misunderstandings (Yildirir and Demirkol, 2018). Three studies pursued further objectives: Samon and Levy (2020) analyzed the universal validity of complex systems, Zarkadis et al. (2017) examined the coherence between mental models, and Chiang and Chiu (2015) evaluated a diagnostic procedure.

Theories on Mental Models in the Selected Studies

The studies analyzed use different theoretical approaches to mental models, often combining different authors and years. A total of 31 relevant authors were identified, with Norman (1983) and Johnson-Laird (1983; 2013) being cited most frequently (Table 10). Other frequently cited works are by Chittleborough (2004), Chittleborough and Treagust (2007), Craik (1943), Wang (2007), Greca and Moreira (2000; 2002), Johnstone (1991; 1993), and Vosniadou and Brewer (1992; 1994). One study (Zarkadis et al., 2017) did not provide any explicit information on the theoretical basis, presumably because this had been dealt with comprehensively in earlier works. The remaining 22 authors, each of whom was cited only once, were grouped together under "Others." It is clear that the majority of the theoretical foundations date from the 1980s to 1990s, which illustrates the historical influence on research into mental models and underscores the importance of this period for further theoretical development.

Table 9: Objectives

Objective	Number of studies
Identification and analysis of mental models	10
Study on the effectiveness of an intervention on mental models	9
Evaluation of a diagnostic and analytical procedure for mental models	1
Testing the universal validity of a theoretical approach in connection with mental models	1
Revealing coherence between different mental models	1

Nature of Mental Models in the Selected Studies

The measurement and representation of mental models is only possible to a limited extent (Gobert and Buckley, 2000; Greca and Moreira, 2000). Nevertheless, the researchers in the selected studies used various methods to capture these structures. The results in Table 11 show that two main approaches can be distinguished: In twelve studies, mental models were derived inductively from consistent patterns of responses, while in ten studies, they were evaluated deductively on the basis of existing theories.

Table 10: Theories about mental models that were presented in the studies

Kodierung	Number of studies
Norman	
1983	3
Johnson-Laird	
1983	2
2013	1
Chittleborough and Treagust	
2007	2
Chittleborough	
2004	1
Craik	
1943	2
Wang	
2007	2
Greca und Moreira	
2000	1
2002	1
Johnstone	
1991	1
1993	1
Vosnaidou and Brewer	
1992	1
1994	1
Not specified	1
Other	22

Table 11: Nature of mental modells

Kodierung	Number of studies
Mental models were derived inductively from consistent patterns in the responses	12
Deductive classification of models according to .	
Park et al. (2009)	3
Kurnaz and Eksi (2015)	2
Albaiti (2017)	1
Jaber und BouJaoude (2012)	1
Chi and Roscoe (2002)	1
Different authors	2

Three studies were based on Park et al. (2009) and classified the models according to correctness and completeness into initial models, intermediate models, and target models. Two other studies used the classification according to Kurnaz and Eksi (2015) into scientific, synthetic, and mental models. Two studies (Samon and Levy, 2020; Zarkadis et al., 2017) combined different theories and authors for their deductive evaluation. It is striking that the majority of the studies used deductive approaches based predominantly on theories and assumptions from the 2000s to 2010s.

Overall, the results illustrate the diversity of methods used to measure and represent mental models. Both inductive and deductive approaches are used, with inductive methods slightly dominating and allowing for flexible recording of individual models.

Impact on Mental Models in the Selected Studies

The results of the selected studies provide insights into the factors influencing mental models. While some studies did not provide explicit information, others identified positive and negative effects (Table 12). Dynamic animations proved to be particularly beneficial, as demonstrated in four studies, including Akaygun (2016a; 2016b). The results show that the use of software such as Kinetic Sketch Pad (Davis et al., 2008) and ChemSense (Schank and Kozma, 2002) had particularly positive effects. Similarly, an improved understanding of the three levels of representation in chemistry was reported, as shown by Albaiti et al. (2022) and Murni et al. (2022), among others. Targeted evaluation and intervention studies with teaching and learning modules, such as those by Praisri and Faikhamta (2020), the use of appropriate materials, a deeper understanding of concepts, and learner motivation also had a positive effect.

Incomplete knowledge and understanding (Putri and Wiyarsi, 2022) and misconceptions (Yildirir and Demirkol, 2018) had a particularly negative effect. Overall, the positive influences predominate, which illustrates that methodological support, motivating learning environments, and targeted didactic measures can make a decisive contribution to promoting mental models.

Table 12: Described influences on mental models from the studies

Influence of mental models	Number of studies
(Dynamic) Animations	4
Understanding the three levels of representation	4
Positive	
Different learning modules	3
Appropriate teaching and learning materials	2
Promoting conceptual understanding	1
Interest and motivation	1
Incomplete knowledge and understanding	3
Negative	
Scientifically unsound ideas	3
Learning by hard	1

Recommendations

A total of seven specific recommendations were identified through inductive coding (Table 13). Nearly half of the studies suggested the use of innovative teaching practices to promote mental models. In particular, the use of software programs, such as Kinetic Sketch Pad (Davis et al., 2008) and ChemSense (Schank and Kozma, 2002), three-dimensional representations, and structured modules was recommended as an effective method.

Another frequently mentioned recommendation was future research in the area of mental models in teaching. Seven studies emphasized the importance of further research and findings to improve the understanding and application of mental models. Five studies highlighted the promotion of the three levels of representation as an important recommendation. This holistic approach to teaching chemical concepts and relationships was considered an effective method for developing mental models in five studies.

The use of misconceptions as learning opportunities and the reorganization of knowledge were each cited as key recommendations by three studies. These approaches emphasize the importance of a learning process in which existing misconceptions are identified and mental models are corrected.

Two studies recommended adapting the framework curricula to specifically support the integration and teaching of mental models.

One study recommended using online assessment to diagnose mental models to efficiently record and evaluate learners' mental models.

DISCUSSION

An analysis of 22 studies on mental models conducted between 2013 and 2022 reveals growing interest in this field, particularly since 2015. A strong geographical focus on Asian countries, especially Indonesia, likely relates to the 2013 curriculum (K-13) implementation (Machali, 2014). Global

Table 13: Recommendations reported in the selected studies

Recommendation	Number of studies
Application of innovative teaching practices (e.g., use of software programs)	10
More research in the field of mental models	7
Promoting the three levels of representation in teaching (e.g., following the Johnstone triangle)	5
Using misunderstandings to design more effective lessons (e.g., Conceptual Change approach)	3
Supporting students in the process of reorganizing knowledge (e.g., Concept Mapping)	3
Revision of the framework curriculum	2
Online assessment system for diagnosing mental models	1

collaborations and international publications highlight the topic's relevance and contribute to research advances and educational reforms.

Methodologically, studies vary widely. More than half were basic studies examining mental models without intervention to identify structure, composition, and misconceptions (Seel, 2014). Intervention studies, for example, using dynamic simulations, examine targeted changes in mental models (Akaygun, 2016a; Akaygun, 2016b). Evaluation studies assess the effectiveness of diagnostic or learning tools (Chiang and Chiu, 2015). Qualitative and mixed-method approaches dominate, particularly semi-structured interviews, which provide deep insights into learners' thinking (Döring and Bortz, 2016; Greca and Moreira, 1997). Drawings and computer-based techniques were used less frequently, though the latter hold potential for supporting mental model development (Seel, 2017). Method choice depends on study objective, model type, and resources, explaining the observed diversity.

Content-wise, studies focus on foundational chemical concepts such as chemical bonds, atomic structures, and acid-base topics (Yildirir and Demirkol, 2018). Basic studies identify misconceptions and provide a foundation for interventions, which in turn allow assessment of validity, coherence, and effectiveness of diagnostic procedures (Samon and Levy, 2020; Zarkadis et al., 2017; Chiang and Chiu, 2015).

Theoretically, studies draw on foundational works from the 1980s to 1990s, especially Norman and Johnson-Laird (Norman, 1983; Johnson-Laird, 1983), complemented by more recent approaches (Gentner and Stevens, 1983). Deductive evaluations enhance comparability, while inductive approaches reveal new patterns. Focusing on central theories could improve the comparability of future deductive studies.

Regarding influences on mental models, positive effects include dynamic animations (Akaygun, 2016a; Akaygun, 2016b), attention to three levels of representation (Albaiti et al., 2022; Murni et al., 2022), targeted teaching materials, motivation, and conceptual understanding. Negative influences, such as incomplete knowledge or misconceptions (Putri and Wiyarsi, 2022; Yildirir and Demirkol, 2018), emphasize the need for careful knowledge transfer and targeted interventions (Buckley, 2000).

Practical implications are clear: Innovative methods such as simulations, software, and three-dimensional representations enhance teaching of the three levels of representation and turn misconceptions into learning opportunities (Seel, 2017). Curriculum and syllabus adjustments are essential to integrate present research findings into education.

In summary, basic, intervention, and evaluation studies complement each other; methodological diversity is necessary for a comprehensive understanding, and targeted educational measures support mental model development. Future research should update theoretical foundations, assess the effectiveness of innovative methods, and explore mechanisms underlying influences on mental models to improve student learning.

OUTLOOK

Research on mental models in teaching continues to offer significant potential for new insights. Further studies are needed to evaluate and enhance the effectiveness of methods and practices for promoting mental models. Innovative teaching approaches – such as structured research-based modules (cf. Murni et al., 2022) or model-based argumentation (cf. Praisri and Faikhamta, 2020) – as well as digital tools are proving particularly promising. They enable both further development of research and a differentiated analysis of the opportunities and risks of technology-supported learning.

To foster learners' understanding and application of mental models, educators and institutions should actively integrate research findings into practice. Curriculum adaptation is essential for systematically translating scientific knowledge into education. The implementation of Indonesia's K-13 curriculum exemplifies how educational reforms can stimulate national research interest in chemistry education (Machali, 2014).

Continuous investigation of factors influencing mental models remains crucial. Identifying negative influences, such as misconceptions, supports targeted interventions and improves knowledge transfer. Future research should focus on elucidating the mechanisms linking these factors to mental models, thereby optimizing teaching and learning and supporting the development of well-founded student mental models.

LIMITATIONS

This literature review has several limitations that should be considered when interpreting the findings. The selection of databases, search terms, and inclusion criteria may have excluded relevant studies. International studies also posed linguistic and contextual challenges, as relevant research may have been published under different terms or in other languages, and variations in study quality may affect the validity of the results. In addition, the 10-year time frame and exclusive focus on chemistry education limit the generalizability of the findings. Extending the review to older or more recent studies, as well as to other disciplines, would be necessary to provide a more comprehensive understanding of mental models in education.

REFERENCES

Akaygun, S. (2016a). Visualizing condensation. In: *Advances in Early Childhood and K-12 Education.* United States: IGI Global, pp. 172-185.

Akaygun, S. (2016b). Is the oxygen atom static or dynamic? The effect of generating animations on students' mental models of atomic structure. *Chemistry Education Research and Practice,* 17(4), 788-807.

Albaiti, A. (2017). *Praktikum Kimia Fisika Berbasis Model Mental untuk Meningkatkan Keterampilan Generik Sains dan Keterampilan Berpikir Kritis Mahasiswa Calon Guru*. Indonesia: Universitas Pendidikan Indonesia.

Albaiti, A., Jukwati., & Lepa, A.A. (2022). Solubility and solubility product phenomena: Papua senior high school students mental model. *Journal of Turkish Science Education,* 19(2), 481-495.

Almadhoun, E., & Parham-Mocello, J. (2021). Exploratory Study on Accuracy of Students' Mental Models of a Singly Linked List. In: *2021 IEEE Frontiers in Education Conference (FIE)*.

Amalia, F.R., Ibnu, S., Widarti, H.R., & Wuni, H. (2018). Students' mental models of acid and base concepts taught using cognitive apprenticeship learning model. *Jurnal Pendidikan IPA Indonesia*, 7(2), 187-192.

Benzer, A.İ., & Ünal, S. (2021). Models and modelling in science education in Turkey: A literature review. *Journal of Baltic Science Education*, 20(3), 344-359.

Buckley, B.C. (2000). Interactive multimedia and model-based learning in biology. *International Journal of Science Education*, 22(9), 895-935.

Cárdenas-Figueroa, A., & Navarro, A.H. (2020). Overview of mental models research using bibliometric indicators. *Cognitive Processing,* 21(2), 155-165.

Chen, S., She, H., & Hsiao, M. (2015). Using Eye-Tracking to Investigate the Different 3D Representation on Students' Mental Model Construction. In: *Conference: 2015 IEEE 15th International Conference on Advanced Learning Technologies (ICALT)*.

Chi, M.T.H., & Roscoe, R.D. (2002). The processes and challenges of conceptual change. In: *Kluwer Academic Publishers eBooks.* Berlin: Springer, pp. 3-27.

Chiang, W.W., & Chiu, M.H. (2015). Using an on-line assessment system to diagnose students' mental models in chemistry education. *Turkish Online Journal of Educational Technology*, 14(1), 163-178.

Chittleborough, G. (2004). *The Role of Teaching Models and Chemical Representations in Developing Students' Mental Models of Chemical Phenomena*. Available from: https://hdl.handle.net/20.500.11937/763 [Last accessed on 2026 Mar 02].

Chittleborough, G., & Treagust, D.F. (2007). The modelling ability of non-major chemistry students and their understanding of the sub-microscopic level. *Chemistry Education Research and Practice*, 8(3), 274-292.

Clarivate. (2023). *EndNote 21* [Computer Software]. Clarivate. Available from: https://endnote.com [Last accessed on 2026 Mar 02].

Clement, J. (2000). Model based learning as a key research area for science education. *International Journal of Science Education*, 22(9), 1041-1053.

Clement, J. (2008). *Creative Model Construction in Scientists and Students: The Role of Imagery, Analogy, and Mental Simulation*. Berlin: Springer Verlag.

Cobb, P., Yackel, E., & McClain, K. (2000). Symbolizing and communicating in mathematics classrooms. In: *Routledge eBooks*. London: Routledge.

Craik, K. (1943). *The Nature of Explanation*. Cambridge, UK: Cambridge University Press.

Davis, R.C., Colwell, B., & Landay, J.A. (2008). *K-Sketch: A "Kinetic" Sketch Pad for Novice Animators*. *Paper Presented at the 26th Annual ACM Conference on Human Factors in Computing Systems (CHI 2008)*. Florence, Italy.

Derkach, T.M. (2021). The origin of misconceptions in inorganic chemistry and their correction by computer modelling. *Journal of Physics Conference Series,* 1840(1), 012012.

Döring, N., & Bortz, J. (2016). *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften*. Berlin: Springer Science+Business Media.

Gates, S. (2002). Review of methodology of quantitative reviews using meta-analysis in ecology. *Journal of Animal Ecology*, 71(4), 547-557.

Gentner, D., & Stevens, A.L. (Eds.). (1983). *Mental Models*. Hillsdale: Lawrence Erlbaum.

Gilbert, J., & Boulter, C.J. (1998). Learning science through models and modelling. In: *Springer eBooks.* London: Kluwer Academic, pp. 53-66.

Glynn, S.M. (1997). Drawing mental models. *The Science Teacher*, 64(1), 30-32.

Gobert, J.D., & Buckley, B.C. (2000). Introduction to model-based teaching and learning in science education. *International Journal of Science Education*, 22(9), 891-894.

Goh, J. (2013). *The 15th International Conference on Biomedical Engineering: ICBME 2013, 4th to 7th December 2013*. Singapore: Springer.

Gomersall, J.S., Jadotte, Y.T., Xue, Y., Lockwood, S., Riddle, D., & Preda, A. (2015). Conducting systematic reviews of economic evaluations. *International Journal of Evidence-Based Healthcare*, 13(3), 170-178.

Greca, I.M., & Moreira, M.A. (1997). The kinds of mental representations--models, propositions and images--used by college physics students regarding the concept of field. *International Journal of Science Education*, 19(6), 711-724.

Greca, I.M., & Moreira, M.A. (2000). Mental models, conceptual models, and modelling. *International Journal of Science Education*, 22(1), 1-11.

Greca, I.M., & Moreira, M.A. (2002). Mental, physical, and mathematical models in the teaching and learning of physics. *Science Education,* 86(1), 106-121.

Halford, G.S. (1993). *Children's Understanding: The Development of Mental Models*. Oxford: Psychology Press.

Halford, G.S., Bain, J.D., Maybery, M.T., & Andrews, G. (1998). Induction of relational schemas: Common processes in reasoning and complex learning. *Cognitive Psychology*, 35(3), 201-245.

Harris, J.D., Quatman, C.E., Manring, M.M., Siston, R.A., & Flanigan, D.C. (2014). How to write a systematic review. *The American Journal of Sports Medicine*, 42(11), 2761-2768.

Harrison, A.G., & Treagust, D.F. (2000). A typology of school science models. *International Journal of Science Education*, 22(9), 1011-1026.

Ifenthaler, D., & Seel, N.M. (2011). A longitudinal perspective on inductive reasoning tasks. Illuminating the probability of change. *Learning and Instruction,* 21(4), 538-549.

Jaber, L.Z., & BouJaoude, S. (2012). A macro-micro-symbolic teaching to promote relational understanding of chemical reactions. *International Journal of Science Education*, 34(7), 973-998.

Johnson-Laird, P.N. (1983). *Mental Models: Towards a Cognitive Science of Language, Inference, and Consciousness*. Cambridge: Cambridge University Press.

Johnson-Laird, P.N. (2004). The history of mental models. In: Manktelow, K., & Chung, M, (Eds.), *Psychology of Reasoning: Theoretical and Historical Perspectives.* London: Taylor and Francis, pp. 179-212.

Johnson-Laird, P.N. (2013). Mental models and cognitive change. *Journal of Cognitive Psychology*, 25(2), 131-138.

Johnstone, A.H. (1991). Why is science difficult to learn? Things are seldom what they seem. *Journal of Computer Assisted Learning*, 7(2), 75-83.

Johnstone, A.H. (1993). The development of chemistry teaching: A changing response to changing demand. *Journal of Chemical Education*, 70(9), 701.

Jonassen, D.H. (2000). *Computers as Mindtools for Schools: Engaging Critical Thinking*. 2nd ed. United States: Prentice Hall/Merrill.

Kiswandari, K., & Ridwan, A. (2020). Analysis of students mental models through POE (predict observe explain) method in salt hydrolisis material. *JTK (Jurnal Tadris Kimiya)*, 5, 80-90.

Kitchenham, B., & Charters, S. (2007). *Guidelines for Performing Systematic Literature Reviews in Software Engineering*. EBSE Technical Report, Software Engineering Group, School of Computer Science and Mathematics, Keele University; Department of Computer Science, University of Durham, 2.3.

Krumm, B., & Scharf, V. (1992). Bildung durch chemieunterricht. Gedanken über die ziele und wege eines zeitgemässen chemieunterrichts. In: Scharf, V., (Eds.), *Zur Professionalisierung der Chemielehrerbildung. Perspektiven und Konsequenzen für die Lehrerbildung Naturwissenschaft und Unterricht.* Vol. 11. Berlin: ResearchGate, pp. 13-22.

Kurdziel, J.P., Beilfuss, M., & Libarkin, J.C. (2003). Research methodologies in science education: Mental models and cognition in education. *Journal of Geoscience Education,* 51(1), 121-126.

Kurnaz, M.A., & Eksi, C. (2015). An analysis of high school students' mental models of solid friction in physics. *Kuram ve Uygulamada Eğitim Bilimleri*, 15, 787-795.

Larson, K., Long, G.R., & Briggs, M.S. (2012). Periodic properties and inquiry: Student mental models observed during a periodic table puzzle activity. *Journal of Chemical Education,* 89(12), 1491-1498.

Machali, I. (2014). Kebijakan perubahan kurikulum 2013 dalam menyongsong Indonesia Emas tahun 2045. *Jurnal Pendidikan Islam,* 3(1), 71.

Mayer, R.E. (1989). Models for understanding. *Review of Educational*

Research, 59(1), 43-64.

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D.G. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *BMJ*, 339(jul21 1), b2535.

Moreira, M.A. (1996). Modelos Mentais (Mental Models). *Investigações em Ensino de Ciências,* 1(3), 193-232.

Murni, H.P., Azhar, M., Ellizar, E., Nizar, U.K., & Guspatni, G. (2022). Three levels of chemical representation-integrated and structured inquiry-based reaction rate module: Its effect on students' mental models. *Journal of Turkish Science Education*, 19(3), 758-772.

Nasir, N., Anuar, M., Yahaya, N., & Ibrahim, N.H. (2019). Assessment of students' mental models regarding hydrogen chloride molecule and its ionization properties among secondary school students. *Man in India*, 13(97), 41-49.

Nitz, S., & Fechner, S. (2018). Mentale modelle. In: *Springer eBooks.* Berlin: Springeer, pp. 69-86.

Norman, D.A. (1983). Some observations on mental models. In: Gentner, D., & Stevens, A.L., (Eds.), *Mental Models*. United States: Lawrence Erlbaum Associates, pp. 7-14.

Okoli, C., & Schabram, K. (2010). A Guide to Conducting a Systematic Literature Review of Information Systems Research. *SSRN Electronic Journal*, 10, 879-910.

Park, E.J., Light, G.J., Swarat, S., & Drane, D. (2009). Understanding Learning Progression in Student Conceptualization of Atomic Structure by Variation Theory for Learning. In: *The Learning Progressions in Science Conference*, Iowa City, Iowa, USA.

Pollock, A., & Berge, E. (2018). How to do a systematic review. *International Journal of Stroke,* 13(2), 138-156.

Praisri, A., & Faikhamta, C. (2020). Enhancing students' mental models of chemical equilibrium through argumentation within model-based learning. *International Journal of Learning, Teaching and Educational Research,* 19(7), 121-142.

Putri, D.N., & Wiyarsi, A. (2022). Identification of mental models of high school students on acid-base theory. In: *Nucleation and Atmospheric Aerosols*. United States: American Institute of Physics.

Rahmi, C.H., Katmiati, S., Wiji., & Mulyani, S. (2017). Students' mental models on the solubility and solubility product concept. *AIP Conference Proceedings*, 1848, 030001.

Ritchie, A.D., & Craik, K.J.W. (1943). *The Nature of Explanation*. Cambridge: Cambridge University Press.

Rumelhart, D.E., Smolensky, P., McClelland, J.L., & Hinton, G.E. (1986). Schemata and sequential thought processes in PDP models. In: *Elsevier eBooks.* Netherlands: Elsevier, pp. 224-249.

Samon, S., & Levy, S.T. (2020). Interactions between reasoning about complex systems and conceptual understanding in learning chemistry. *Journal of Research in Science Teaching,* 57(1), 58-86.

Schank, P., & Kozma, R. (2002). Learning chemistry through the use of a representation-based knowledge-building environment. *Journal of Computers in Mathematics and Science Teaching,* 21(3), 253-279.

Seel, N.M. (1999). Educational diagnosis of mental models: Assessment problems and technology-based solutions. *Journal of Structural Learning and Intelligent Systems,* 14(2), 153-185.

Seel, N.M. (2014). Model-based learning and performance. In: *Springer eBooks.* Berlin: Springer, pp. 465-484.

Seel, N.M. (2017). Model-based learning: A synthesis of theory and research. *Educational Technology Research and Development*, 65(4), 931-966.

Sunyono, S., & Sudjarwo, S. (2018). Mental models of atomic structure concepts of 11th-grade chemistry students. *Asia-Pacific Forum on Science Learning and Teaching,* 19(1), 9.

Tasker, R. (2016). ConfChem conference on interactive visualizations for chemistry teaching and learning: Research into practice-visualizing the molecular world for a deep understanding of chemistry. *Journal of Chemical Education*, 93(6), 1152-1153.

Thagard, P. (2010). How brains make mental models. In: *Studies in Computational Intelligence*. Berlin: Springer Nature, pp. 447-461.

Tümay, H. (2014). Prospective chemistry teachers' mental models of vapor pressure. *Chemistry Education Research and Practice,* 15(3), 366-379.

Vosniadou, S., & Brewer, W.D. (1992). Mental models of the earth: A study of conceptual change in childhood. *Cognitive Psychology*, 24(4), 535-585.

Vosniadou, S., & Brewer, W.D. (1994). Mental models of the day/night cycle. *Cognitive Science,* 18(1), 123-183.

Wang, C. (2007). *The Role of Mental-Modeling Ability, Content Knowledge, and Mental Models in General Chemistry Students' Understanding about Molecular Polarity*; [Thesis].

Wardah, A.C., & Wiyarsi, A. (2020). A systematic review: How are mental models of chemistry concepts? *Universal Journal of Educational Research,* 8(2), 332-345.

Xiao, Y., & Watson, M.E. (2017). Guidance on conducting a systematic literature review. *Journal of Planning Education and Research,* 39(1), 93-112.

Yildirir, H.E., & Demirkol, H.G. (2018). Identifying mental models of students for physical and chemical change. *Journal of Baltic Science Education,* 17(6), 986-1004.

Zarkadis, N., Papageorgiou, G.Z., & Stamovlasis, D. (2017). Studying the consistency between and within the student mental models for atomic structure. *Chemistry Education Research and Practice*, 18(4), 893-902.