Crucial Cognitive Skills in Science Education: A Systematic Review

(Received 8 January 2020; Revised 16 May 2020; Accepted 16 May 2020)

Uswatun Hasanah¹, Kinya Shimizu²

¹,²Graduate School for International Development and Cooperation, Hiroshima University, Hiroshima, Japan

Corresponding Author: *uswatunhasanah216@yahoo.com

DOI: 10.30870/jppi.v6i1.7140

Abstract

This systematic review focuses on identifying three common cognitive skills in science education—process skills, critical thinking skills, and reasoning skills—in order to find the crucial skills in science education. The inclusion and exclusion criteria were created. In total, 78 articles published in 17 countries, namely USA, Turkey, Indonesia, Malaysia, Iran, Palestine, Thailand, Nigeria, Jamaica, Israel, Kenya, Oman, Columbia, China, Philippines, Korea, Canada, were selected. The reviewed studies were published from 1998 to 2019. Fifty-seven studies were reported as journal publications and 21 studies came from conference proceedings. The results indicate that crucial skills exist such as science process skills (inference, measuring, identifying and controlling variable, definition operational variable, and explanation), critical thinking skills (interpreting data, inference, and evaluation), and reasoning skills (all subskills), and also revealed the relationship among them. This study concludes that the crucial skills in science education are mostly located in the reasoning skills domain.

Keywords: Cognitive Skills, Science Process Skills, Critical Thinking Skills, Reasoning Skills, Science Education
INTRODUCTION

Cognitive skills such as science process, critical thinking, and reasoning are common and most prominent to social scientists, teachers, and all facilitators in science education. Many psychologists consider cognitive skills have a close relationship with students’ overall capacity to learn (Han, 2013). Cognitive skills can support the student in constructing knowledge, assumptions, competence, and the ability to solve problems and formulate results. Additionally, some researchers have stated that cognitive skills are connected to each other, either directly and/or indirectly (Lawson, 1995; Ozgelen, 2012). However, the variety of skill labels are confusing for teachers who are required to develop the more than necessary skills of students (Bailin, 2002; Lewis & Smith, 1993; Niu et al., 2013; Zimmerman, 2000).

Numerous studies have tried to identify improvement, particularly in science process skills (Huppert, Lomask, & Lazarowitz, 2002; Lati, Supasorn, & Promarak, 2012; Ongowo & Indoshi, 2013), critical thinking skills (Duran & Dokme, 2016; Shin, 1998; Zhou, Huang, & Tian, 2013), and reasoning skills (Mendoza et al., 2018; Remigio et al., 2014), and these studies have confirmed the importance of each. However, with regard to the development of cognition domain in students, it remains unclear whether these skills are related. Additionally, the literature reviewed in preparation for this paper suggests that little or no research has been conducted to identify the most crucial cognitive skills in science education. Therefore, this systematic review is focused on identifying the most crucial skill in science education among the common cognitive skills. The author intends to respond to the following research questions:

1. What are the crucial skills in science process skills, critical thinking skills, and reasoning skills?
2. Where are the crucial skills located in the cognition stages?

Cognitive Skills

Science process skills (SPS) are mental abilities that can be practiced, learned, and developed by children through the learning process, making the students better able to meet the challenges of the 21st century (Balfakih, 2010; Osman & Vebrianto, 2013). SPS are essential to teaching the ways of reaching knowledge and they can ensure that students have a meaningful learning experience (Rauf et al., 2013). Today, the phrase “science process skills” is commonly used and based on Science-A Process Approach (SAPA); these skills can be classified into two categories: basic and integrated SPS. Germann and
Aram (1996) in Rauf et al. (2013) defined basic skills as the intellectual foundation of scientific inquiry. Basic skills are the preconditions to integrated process skills, which are the terminal skills for problem-solving or conducting science experiments. The sub-skills of SPS are observation, inferring, measuring, communicating, classifying, predicting, controlling variables, defining operationally, formulating hypotheses and models, interpreting data, and experimenting (Padilla, 1990).

Furthermore, there are widely contrasting views of critical thinking skills (CTS). Some of the views highlight the range of perspectives developed around the aspect of education. In summary, CTS is defined as the mental act of reviewing, evaluating, or appraising something in an attempt to make judgments, inferences or meaning about it in a rational, reasoned way (McGroger, 2007). CTS is considered to be intellectually engaged, skillful, and responsible thinking. It facilitates good judgment that requires the application of assumptions, knowledge, competence, and the ability to challenge one’s thinking, as well as self-correction, monitoring the reasonableness of thinking, and reflexivity. One characteristic that uniquely defines CTS is that individuals are capable of stepping back and reflecting on the quality of their thinking (Niu et al., 2013). In this research, the authors adapted the core idea of CTS from Facione (1990), who provided in much more detail the descriptors of the associated characteristics. The subskills of CTS are interpretation, analysis, evaluation, inference, explanation, and self-regulation.

The last cognitive skill is reasoning skills (RS) from Lawson (2000), who developed the theory based on psychologists’ theory of cognitive development in the last two stages—empirical-inductive (EI) thought and hypothetical-deductive (HD) thought. EI thinking patterns enable the child to order accurately and describe perceptible objects, events, and situations within their world. In this stage, the child starts using language for their logical reasoning. Conservation was taken as one of the subskills in RS. Meanwhile, HD thinking patterns allow the adolescent to go beyond descriptions and create and test hypothetical explanations for what is encountered (Lawson, 1995). The subskills in RS, based on Lawson (2000), are conservation law (EI), proportional reasoning (HD1), identification and control of variables (HD2), probabilistic reasoning (HD3), correlational
reasoning (HD4), and hypothetical-deductive reasoning (HD5).

**Relationships among Cognitive Skills**

As mentioned above, the theory of cognitive development has been established by many psychologists. Piaget (1966) investigated cognitive development in terms of how the child perceives the environment and the world based on observation and interview. According to Piaget’s theory, cognitive development can be divided into four stages, based on age. This systematic review focuses on the last two stages (concrete operational and formal operational) that were previously introduced as EI and HD. Concrete reasoning begins from seven or eight years of age and applies a new level, such as naming, describing, and classifying; formal operational reasoning begins from adolescence and older. In this stage, some children become increasingly capable of using language to apply the deductive pattern of thinking to hypothetical rather than empirical representations. The epistemology of the concrete reasoning stage thinker is one of observation: What causes events? In order to find an answer to this question, the child (or the inquirer) would have to observe some event(s). The epistemology of the formal reasoning stage thinker is vastly different: What causes events? To find the answer, one must first mentally create several possible causes, deduce their potential consequences, and then observe the results of experimental manipulations to support or reject the possibilities (Lawson, 1995). The term formal RS is typically used by researchers to define more complex skills such as SPS and CTS. For the purposes of this study, the details of the relationship among SPS, CTS, and RS are illustrated in Figure 1.

![Conceptual framework of relationships among cognitive skills](image-url)

Figure 1. Conceptual framework of relationships among cognitive skills
This model involves five groups, based on the skill’s pattern, as listed below:
1. Observation, measuring, communicating, classifying, defining operational variable;
2. Identification and controlling variable (HD2), predicting (HD3), formulating hypothesis (HD5), experimental design (HD5), conducting experiment (HD5);
3. Interpreting data (HD1), inference (HD4), analysis (HD4);
4. Simple explanation (conservation/EI), explanation (HD5), evaluation (HD5);
5. Self-regulation.

The first group (observation, measuring, communicating, classifying, defining the operational variable) is characteristic of the basic SPS (Padilla, 1990). This group is a part of the empirical thinking stage but does not include RS or CTS. The second group paints the relationship between SPS and RS, which cover the identification and controlling variable (HD2) and predicting skills (HD3) (Han, 2013), followed by formulating hypothesis, experimental design and conducting experiment that are covered under hypothetical deductive reasoning (HD5). Ozgelen (2012) revealed that the term formal RS is typically used by researchers to define more complex skills and integrated SPS.

The third group talks about the relationship among SPS, CTS, and RS consisting of two skills: (i) Interpreting data (HD1) and (ii) inference (HD4) and analysis (HD4). The fourth group states the relationship between CTS and RS in terms of explanation skills which is divided into two: simple explanation (EI) and explanation (HD5) and evaluation (HD5). The fifth group is self-regulation that helps the student arrange a strategy to find the solution. This skill includes the CTS group but not the other two cognitive skills.

**METHOD**

A systematic literature review identifies, selects, synthesizes, and appraises the studies that meet the prespecified inclusion criteria for investigating the status of current research, based on research objectives (Knoll et al., 2018). This study consisted of several processes: (a) creating the detail of criteria for inclusion and exclusion empirical studies, (b) conducting literature search in electronic databases, based on the inclusion criteria, (c) finalizing the literature research and reading the details of each study, (d) identifying the pattern of the studies, and (e) synthesizing the pattern in order to answer the research questions.
Inclusion Criteria

To conduct the systematic review, the following inclusion and exclusion criteria were established: (a) The studies should conduct empirical research for both qualitative and quantitative, (b) the studies should focus on improving cognitive skills in science education (SPS, CTS, and RS), (c) the skill dimension for each cognitive skill should be based on the experts Michael J. Padilla for science process, Ils, Peter Facione for CTS, and Anton E. Lawson for RS, or at least clearly explain the subskill’s definition, (d) the studies results need to clearly mention the detailed progress of each subskill in order to fit the objectives and answer the research questions of this systematic review, (e) the sample of the studies should include students from primary to university levels, (f) the studies should be published in full-text journals as a journal article or conference proceeding, (g) the studies should be published in English and Indonesian, and (h) there was no time restriction for the studies.

Literature Search and Analysis

The literature search on electronic databases was conducted through Education Resources Information Center (ERIC), Scopus, Web of Science and Google Scholar using every conceivable variety of keywords such as “Cognitive”, “Process skill”, Critical thinking skill”, “Reasoning skill”, “Science education”, and “Students” without time restriction. At first identification using the mentioned electronic databases, 696 potential articles were found, as shown in Figure 2.

Afterwards, the authors conducted title and abstract screening. It showed 288 articles eligible for full-text screening, and it finally resulted in 78 articles that were ready to be reviewed—32 articles for SPS, 34 articles for CTS, and 12 articles for RS.

In analyzing the data process, the researchers conducting data extraction outlined an overview with characteristics including (a) author and sample size, (b) country, (c) institution, (d) design (measurement and instrument), and (e) findings. The approach adopted for data analysis and reporting was a narrative content analysis based on the expert recommendation from Knoll et al. (2018), to avoid the studies being too heterogeneous in terms of design or the outcome analyzed (Popay et al., 2006). To address the first research question, the findings from the selected studies in this systematic review were divided into two categories—good and crucial. The good category covers all the subskills that obtain a high mean score or the greatest improvement during the
intervention, or a high rate or frequency for being an essential subskill from the study participants. Meanwhile, the crucial category includes the crucial subskills that obtain the lowest mean scores or improvement (decreasing), or a low rate or frequency. For the second research question, the authors analyzed whether all crucial subskills in one cognitive skill can also become the crucial subskills for the other cognitive skills, and how they are related, based on the model we created in Figure 1.

RESULTS AND DISCUSSION

Based on what has been mentioned in the introduction, this systematic review focused on identifying three cognitive skills—SPS, CTS, and RS. In total, 78 articles were selected from 17 countries, namely the USA, Turkey, Indonesia, Malaysia, Iran, Palestine, Thailand, Nigeria, Jamaica, Israel, Kenya, Oman, Columbia, China, Philippines, Korea, and Canada. The reviewed studies were published from 1998 to 2019, 57 of which were journal publications and 21 conference proceedings. In the following paragraphs, the author specifies each skill as a finding in this systematic review.
Crucial Domains in Science Process Skills

Under SPS, 32 studies were reviewed with the characteristics (author and sample size, country, institution, design, measurement and instrument, and findings), showed in Table 1.

Table 1. The Characteristics of 32 Included Studies under SPS

<table>
<thead>
<tr>
<th>Author (Sample Size)</th>
<th>Country</th>
<th>Institution</th>
<th>Design (Measurement Instrument)</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ozturk et al. (2010) (n = 828)</td>
<td>Turkey</td>
<td>Eskisehir Osmangazi University</td>
<td>Experimental, Questionnaire, Grade 7th, secondary school</td>
<td>Recording data skill, observation, classification, measure, and data explication, formulating a hypothesis, modelling, decision skill achieved at a high level</td>
</tr>
<tr>
<td>Zeidan &amp; Jayosi (2015) (n = 159)</td>
<td>Palestine</td>
<td>Al-Quds University</td>
<td>Experimental, Questionnaire, Grade 10th, secondary school</td>
<td>Observation skill, predicting, measuring, data interpreting, communicating, hypothesizing skill get high rank</td>
</tr>
<tr>
<td>Yilmaz (2019) (n = 332)</td>
<td>Turkey</td>
<td>Karamanoglu Mehmetbey University</td>
<td>Experimental, Questionnaire, Grade 3rd, 4th, primary school</td>
<td>Classification skill, controlling variable and experimentation skill get low rank</td>
</tr>
<tr>
<td>Kamba et al. (2018) (n = 203)</td>
<td>Nigeria</td>
<td>Kebbi State University of Science and Technology</td>
<td>Experimental, Questionnaire, Secondary school</td>
<td>Observation, classification, communication skills got a high level</td>
</tr>
<tr>
<td>Duruk et al. (2017) (n = 307)</td>
<td>Turkey</td>
<td>Adiyaman University</td>
<td>Survey, Method and document analysis, Grade 5th, 6th, 7th, 8th,</td>
<td>Communicating skills, classifying, controlling variables and hypothesising, and experimentation skills got a low rank.</td>
</tr>
</tbody>
</table>

Inference, measurement, and prediction skills got a low level.

Observation skill, measuring, predicting, and data interpreting skill got high rank.

Communicating skills, classifying, controlling variables and hypothesising, and experimentation skills got a low rank.

Observation skill, communicating, classifying, interpreting data, experimenting, and modelling skills was the most represented.

Inferring, measuring, predicting, controlling.
<table>
<thead>
<tr>
<th>Author (Sample Size)</th>
<th>Country</th>
<th>Institution</th>
<th>Design (Measurement Instrument)</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tekerci and Kandir (2017) (n = 40)</td>
<td>Turkey</td>
<td>-</td>
<td>• Quasi-experimental • Questionnaire • Preschool and Nursery classes</td>
<td>variables, defining operationally, formulating hypothesis skills scored least. Observation, comparison, classification, measurement, communication, inference, predicting skill showed statistically significant difference.</td>
</tr>
<tr>
<td>Ting and Siew (2014) (n = 119)</td>
<td>Malaysia</td>
<td>Universiti Malaysia Sabah</td>
<td>• Quasi-Experimental • Questionnaire • Grade 5th, primary school</td>
<td>Observing skill, communicating, and classifying skill showed the greatest improvement.</td>
</tr>
<tr>
<td>Sahhyar and Febriani (2017) (n = 62)</td>
<td>Indonesia</td>
<td>State University of Medan</td>
<td>• Quasi-Experimental • Observation sheet • Grade 11th, secondary school</td>
<td>Inferring skill, predicting, and controlling variables skill have the least increment. Observing skill, questioning, interpreting, classifying, predicting, communicating, planning, applying concept, generalizing skill had the highest percentage average.</td>
</tr>
<tr>
<td>Gultep and Kilic (2015) (n = 34)</td>
<td>Turkey</td>
<td>Dumlupinar University</td>
<td>• Quasi-experimental • Questionnaire • Grade 11th, secondary school</td>
<td>Inferring skill and making hypothesis skills had the lowest percentage average. Designing experiments skill had no statistically significant difference. Forming data table skill, graph drawing, graph interpretation, determining the variables and building up a hypothesis, changing, and controlling variables skill got the significant effect of the treatment.</td>
</tr>
<tr>
<td>Harahap et al. (2019) (n = 94)</td>
<td>Indonesia</td>
<td>State University of Medan</td>
<td>• Quasi-Experimental • Questionnaire • Biology program Faculty of Mathematics and Science, University level</td>
<td>All results in all indicators of SPS showed significant differences among students. Exceptions were for asking questions skill, planning an experiment, and implementing concept skills.</td>
</tr>
<tr>
<td>Author (Sample Size)</td>
<td>Country</td>
<td>Institution</td>
<td>Design (Measurement Instrument)</td>
<td>Findings</td>
</tr>
<tr>
<td>----------------------</td>
<td>---------</td>
<td>-------------</td>
<td>----------------------------------</td>
<td>----------</td>
</tr>
</tbody>
</table>
| Aydogdu (2017) (n = 1272) | Turkey | Afyon Kocatepe University | • Experimental  
• Questionnaire  
• Primary school | Primary students gained the highest success percentage in prediction skill, classification, observation.  
The lowest success percentage showed in communicating skill, measurement, and inference skill.  
Students obtained the highest mean rank in formulating hypothesis skill, investigating, inferring, interpreting skill. |
| Irwanto et al. (2019) (n = 43) | Indonesia | Yogyakarta State University | • Quasi-experimental  
• Questionnaire  
• University level |  
The lowest in communicating skill, measuring, experimenting, identifying and controlling variables, and observing skill.  
The subjects’ mean score was low and unsatisfactory; their performance in decreasing order was: interpreting data, recording data, generalizing, formulating hypotheses, and identifying variables. |
| Beaumont-Walters and Soyibo (2010) (n = 305) | Jamaica | - | • Experimental  
• Questionnaire  
• Grade 10th, secondary school | Identifying experimental question, designing investigation, graph data skill was a statistically significant difference in comparing both groups.  
Formulating hypothesis skill was no statistically significant difference in comparing both groups.  
Identifying variables skill, defining operationally, stating hypothesis skill was in the highest rank of performance. |
| Turpin and Cage (2004) (n = 531) | Louisiana | Louisiana Department of Education | • Quasi-Experimental  
• Questionnaire  
• Grade 7th, secondary school |  |
| Ogan-Bekiroglu and Arslan (2014) (n = 17) | Turkey | Marmara University | • True experimental  
• Questionnaire  
• Pre-service physics teacher |  |
| Wahyu et al. (2017) | Indonesia | Jember University | • Pre-experimental |  |

Jurnal Penelitian dan Pembelajaran IPA  
Vol. 6, No. 1, 2020, p. 36-72  
Hasanah & Shimizu
<table>
<thead>
<tr>
<th>Author(s)</th>
<th>Sample Size</th>
<th>Country</th>
<th>Institution</th>
<th>Design (Measurement Instrument)</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ates (2004)</td>
<td>(n = 103)</td>
<td>Turkey</td>
<td>Bolu Abant Izzet Baysal University</td>
<td>Experimental, Questionnaire, Junior college</td>
<td>Formulating question skill, formulating problems, formulating conclusions, classifying and analyzing data, apply concept, and making predictions skills were in less category.</td>
</tr>
<tr>
<td>Saribas and Bayram (2009)</td>
<td>(n = 54)</td>
<td>Turkey</td>
<td>Marmara University</td>
<td>Quasi-experimental, Questionnaire, University level</td>
<td>Defining operationally, interpreting and graphing data skill was a statistically significant difference between transitional and concrete reasoners.</td>
</tr>
<tr>
<td>Mutlu and Temiz (2013)</td>
<td>(n = 496)</td>
<td>Turkey</td>
<td>Nigde University</td>
<td>Experimental, Questionnaire, Secondary school</td>
<td>Identifying &amp; controlling variables skill and Stating hypothesis skills were not statistically significant differences between both reasoners.</td>
</tr>
<tr>
<td>Osman and Vebrianto (2013)</td>
<td>(n = 96)</td>
<td>Malaysia</td>
<td>The National University of Malaysia</td>
<td>Quasi-experimental, Questionnaire, Secondary school</td>
<td>Identifying variables, operationally defining, designing investigations were statistically significant difference means can be improved easier comparing to identifying and stating hypotheses skill as well as graphing and interpreting data. The variations are observed to be statistically meaningful in terms of responding variable identification, controlled variable identification, formulating a hypothesis, variable modification, and control skill. Otherwise, identifying manipulated variable skills and interpreting data skills was not statistically meaningful. A significant difference between groups in classifying skill, predicting, and inference. However, there are no significant differences in observing and communication</td>
</tr>
<tr>
<td>Author (Sample Size)</td>
<td>Country</td>
<td>Institution</td>
<td>Design (Measurement Instrument)</td>
<td>Findings</td>
<td></td>
</tr>
<tr>
<td>----------------------</td>
<td>---------</td>
<td>-------------</td>
<td>----------------------------------</td>
<td>----------</td>
<td></td>
</tr>
</tbody>
</table>
| Jeentho ng et al. (2013) (n = 73) | Thailand | Mahidol University | • Quasi-experimental  
• Questionnaire  
• Grade 11th, Secondary school | Collect data skill, and design experiment skill got the higher mean scores.  
Identifying variables skill and pose question and hypothesis skill got the lowest mean score. |
| Siahaan et al. (2017) (n = 23) | Indonesia | Indonesia University of Education | • Pre-experimental  
• Questionnaire  
• Grade 7th, secondary school | Only predicting skill was in the high criteria;  
Observing skill, summarizing, communicating, and classifying skills was in moderate criteria.  
From three grade, the result shows that predicting skill, formulating a hypothesis, and classifying skill were the highest represented skills. |
| Delen and Keserci oglu (2012) (n = 290) | Turkey | Michigan State University | • Experimental  
• Questionnaire  
• Grades 6th, 7th, 8th | Observing skill, Interpreting data, inferring, defining operationally, and experimenting skills were the lowest represented skills.  
Observing skill, inferring, communicating, interpreting data, experimenting skill was the rated highest frequency.  
Measuring skill, classifying, predicting, controlling variables, defining operationally, formulating hypothesis skills, and formulating model skills were rated the lowest.  
Only identifying and controlling variables skill was identified as “excellent.”  
Defining operationally skill, formulating a hypothesis, experimenting, and interpreting data and drawing conclusion skills were identified as “good and fair.” |
| Ongowo and Indoshi (2013) (n = 10) | Kenya | Maseno University | • Observation  
• School records  
• Secondary school | |
| Lati et al. (2012) (n = 63) | Thailand | Ubon Ratchathani University | • Experimental  
• Questionnaire  
• Grade 11th, secondary school | |
<table>
<thead>
<tr>
<th>Author (Sample Size)</th>
<th>Country</th>
<th>Institution</th>
<th>Design (Measurement Instrument)</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Akinbobola and Afolabi (2010) (n = 10)</td>
<td>Nigeria</td>
<td>University of Uyo</td>
<td>• Observation • School records • Secondary school</td>
<td>Observing skill, calculating, recording, communicating. Manipulating skill was rated highest; measuring skill, comparing, contrasting, drawing, experimenting and investigating, graphing, interpreting, deducing, and formulating model skills were rated lowest.</td>
</tr>
<tr>
<td>Rauf et al. (2013) (n = 24)</td>
<td>Malaysia</td>
<td>Universiti Malaya</td>
<td>• Experimental • Questionnaire • Grade 8th, secondary school</td>
<td>Observing skill, communicating, and experimenting skill was the highest frequency and percentage that inculcate in the lesson. Classification skill, measurement and use of the number, making inference, making a prediction, interpreting data, defining operationally, controlling variables, and forming hypothesis skills were the lowest frequency and percentage. Measurement skill and graph communication skill were a statistically significant difference in two cognitive stages.</td>
</tr>
<tr>
<td>Huppert et al. (2002) (n = 181)</td>
<td>Israel</td>
<td>University of Haifa Tivon</td>
<td>• Experimental • Questionnaire • Grade 10th, secondary school</td>
<td>Classification skill, interpreting data, prediction, evaluating hypothesis, controlling variables, selecting useful data, and designing an experiment skill were not significant in two or all cognitive stages. Indicator percentage of observing skill, planning experiment, classifying, organizing data in the table, and identifying variable skill was higher than inference and communicating skill.</td>
</tr>
<tr>
<td>Laksono et al. (2018) (n = 61)</td>
<td>Indonesia</td>
<td>Yogyakarta State University</td>
<td>• Observation • Observation sheet • Grade 10th, Secondary school</td>
<td></td>
</tr>
<tr>
<td>Author (Sample Size)</td>
<td>Country</td>
<td>Institution</td>
<td>Design (Measurement Instrument)</td>
<td>Findings</td>
</tr>
<tr>
<td>----------------------</td>
<td>---------</td>
<td>-------------</td>
<td>---------------------------------</td>
<td>----------</td>
</tr>
</tbody>
</table>
| Maison et al. (2019) (n = 130) | Indonesia | Jambi University | • Correlational research  
- Observation sheet  
- University level | Overall basic science process skills of physical education students of Jambi University are still considered not good. |
| Prihatnawati, et al. (2017) (n = 138) | Indonesia | State University of Malang | • Quasi-experimental  
- Questionnaire  
- Grade 8th, secondary school | Observing skill and conducting experiment skills got the highest average. |
| Molefe et al. (2016) (n = 75) | South Africa | University of KwaZulu-Natal | • Quantitative research with a qualitative component  
- Questionnaire  
- University level | Observing skill, interpreting data skill, classifying skill, formulating a hypothesis, interpreting data, and experimenting skills are chosen as the most important by the participant. Inferring skill, measuring skill, communicating, and predicting skills are chosen as the least important in science process skills. |

![Percentage of Included Studies](image)

Figure 3. Summary of the findings of 32 included studies in SPS
In terms of first characteristic (the total sample), 6248 participants were examined in these studies that four studies are at the primary level, 20 studies are at the secondary level, and eight studies are at the university level. Secondly, the studies selected mostly comes from Turkey and Asian countries such as Indonesia, Malaysia, and Thailand. Thirdly, twenty-eight studies were conducted quantitatively experimental, pre-, true, and quasi-experimental design. The last characteristic is the finding of the studies that are summarized in Figure 3. It shown that observing skills, communicating skills, measuring, classifying skills, predicting, and making models are included in Good category. Meanwhile, it was also found that there are six sub-skills which emerged as a crucial skill. It is starting from Inference skill with 70.6% of 17 studies that conclude inference as the crucial domain in science process skills (Aydogdu, 2017; Delen & Kesercioglu, 2012; Molefe et al., 2016; Ozturk et al., 2010; Ting & Siew, 2014; Yilmaz, 2019).

Lastly, the experimenting skill appeared in 61.9% of 21 study findings and was put under the crucial category (Gultepe & Kilic, 2015; Kamba et al., 2018; Ogan-Bekiroglu & Arslan, 2014). The studies concluded that students in science education need sufficient physical experiences in order to improve their SPS. Duruk et al. (2017) revealed that these skills pose problems in terms of the science curriculum. It may be expressed that the broken and tricky parts of science handling abilities are influenced by the common structure of science educational modules, reflected in course substance, lesson plans, learning action, and results. Aydogdu (2017) also stated that teachers should develop students' skills in inference and
measuring by requiring the active use of these skills in the classroom.

The participants in these studies seemed to have a problem in designing and conducting the experiment, it was rather difficult for them to pose questions and hypotheses. Various issues such as students’ prior knowledge, learning style, learning process, number of students in the class, time limitation must be considered for successful implementation (Jeenthong et al., 2013). In summary, the crucial subskills in SPS are mostly in the integrated domain with inference as an exception. Furthermore, the studies revealed that the curriculum has an impact on the representation of SPS, and changing it effects the representation of SPS.

**Crucial Domains in Critical Thinking Skills**

In this domain, 34 articles were selected to be reviewed, as listed in Table 2. These consisted 15 paper proceedings and 19 journal articles. The total sample from the selected studies was 3608 participants, who came from many levels of primary, secondary, and university. Four studies were conducted qualitatively, and 30 studies were conducted quantitatively, such as quasi-experiment and true experiment. Most of the selected studies were conducted in Asian countries, namely Malaysia, Indonesia, Thailand, Philippines, China, Korea, and Thailand, two studies were conducted in the USA, Turkey, Iran, and Oman.

**Table 2. The Characteristics of 34 Included Studies in Critical thinking Skills.**

<table>
<thead>
<tr>
<th>Author (Sample Size)</th>
<th>Country</th>
<th>Institution</th>
<th>Design (Measurement Instrument)</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bagheri and Nowrozi (2015)</td>
<td>Iran</td>
<td>Payame Noor University</td>
<td>• Comparative</td>
<td>Evaluation skill and induction skill were the most averages among students.</td>
</tr>
<tr>
<td>(n = 60)</td>
<td></td>
<td></td>
<td>• Questionnaire</td>
<td>Deduction, explanation, and analysis skills were in the lowest average.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Vocational university</td>
<td>Interpretation, evaluation, and self-regulation skills emerged as the highest mean score.</td>
</tr>
<tr>
<td>Dilekli (2017)</td>
<td>Turkey</td>
<td>Aksarary University</td>
<td>• Experimental</td>
<td></td>
</tr>
<tr>
<td>(n = 225)</td>
<td></td>
<td></td>
<td>• Questionnaire</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Grade 5th to 8th</td>
<td></td>
</tr>
<tr>
<td>Author (Sample Size)</td>
<td>Country</td>
<td>Institution</td>
<td>Design (Measurement Instrument)</td>
<td>Findings</td>
</tr>
<tr>
<td>----------------------</td>
<td>---------</td>
<td>-------------</td>
<td>----------------------------------</td>
<td>----------</td>
</tr>
</tbody>
</table>
| Kumar (2017) (n = 214) | Oman | Nizwa College of Technology | • Experimental  
• Questionnaire  
• College students | Analysis, inference, and explanation skills emerged as the lowest mean score. Assumption, deductions, and arguments skill were the highest mean score after study. Interpretations skill and inference skill were the lowest mean score after study. |
| Siriwat and Katwibun (2017) (n = 47) | Thailand | Chiang Mai University | • Experimental  
• Questionnaire  
• Grade 11th, secondary school | Explanation issues, and evidence were the highest rated. Influence of context and assumptions, student’s position and conclusions and outcomes were the lowest rated |
| Putra and Prayitno (2018) (n = 188) | Indonesia | Sebelas Maret University of Surakarta | • Quasi experimental  
• Questionnaire  
• Grade 11th, secondary school | Interpretation skill, analysis, and explanation skill had the highest percentage. Evaluation skill, self-regulation and concluding skill had the lowest percentage. |
| Usmeldi, Amini, and Trisna (2017) | Indonesia | State University of Padang | • Experimental  
• Questionnaire  
• Secondary school | Analysis and induction skill had the highest percentage; inference, evaluation, and deduction skill |
<table>
<thead>
<tr>
<th>Author (Sample Size)</th>
<th>Country</th>
<th>Institution</th>
<th>Design (Measurement Instrument)</th>
<th>Findings</th>
</tr>
</thead>
</table>
| Ramos et al. (2013)  | Philippines | Benguet State University | • Experimental  
• Questionnaire, Observation sheet  
• University school | had the lowest percentage. Analysis, comparison, inference and evaluation skills were in the average level and below average level. |
| (n = 393)            |         |             |                                  |          |
| Kong (2014)          | Hong Kong | The Hong Kong Institute of Education | • Experimental  
• Questionnaire  
• Grade 11th, secondary school | Hypothesis identification, induction, and deduction skill had the highest mean score. Evaluation skill and explanation skill had the lowest mean score. |
| (n = 107)            |         |             |                                  |          |
| Duran and Dokme (2016)| Turkey | Giresun University | • Experimental  
• Questionnaire  
• Grade 6th, primary school | In this study, the result shows a significant difference between both groups in terms of the measured analysis skill, evaluation, inference, interpretation, explanation, and self-regulation. |
| (n = 90)             |         |             |                                  |          |
| Zhou et al. (2013)   | China   | Normal University | • Quasi experimental  
• Questionnaire  
• Grade 12th, secondary school | Analysis skill was statistically significant different in both groups, but evaluation skill and inference skill were not. |
| (n = 119)            |         |             |                                  |          |
| Hairida (2016)       | Indonesia | University of Tanjungpura Pontianak | • Quasi experimental  
• Questionnaire  
• Grade 7th, secondary school | Analysis skill and explanation skills had shown the highest average score. Interpretation |
|                      |         |             |                                  |          |

Jurnal Penelitian dan Pembelajaran IPA  
Vol. 6, No. 1, 2020, p. 36-72
<table>
<thead>
<tr>
<th>Author (Sample Size)</th>
<th>Country</th>
<th>Institution</th>
<th>Design (Measurement Instrument)</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shin (1998) (n = 234)</td>
<td>Korea</td>
<td>Ewha Womans University</td>
<td>Experimental, Questionnaire, Secondary school</td>
<td>The skill, inference, and evaluation skill had shown the lowest average. Interpretation, analysis, and inference skill were statistically significant difference in both groups. Evaluation skill and deduction skill were not significant in both groups. Inference skill and evaluation skill got the highest mean score compared to interpretation, analysis and explanation. Inference skill, Evaluation, and explanation skill got the highest mean score compared to interpretation, analysis. The result of the analysis indicated that 58.3% of the students got the proficient level achievement, spreading from the critical skills of communication, analysis (interpretation) and synthesis. It was found to perform well in classification, but</td>
</tr>
<tr>
<td>Asefi and Imani, (2018) (n = )</td>
<td>Iran</td>
<td>Tabriz Islamic Art University</td>
<td>Quasi experimental, Questionnaire, University level</td>
<td></td>
</tr>
<tr>
<td>Stephenson et al. (2019) (n = 159)</td>
<td>USA</td>
<td>Florida International University</td>
<td>Quasi experimental, Questionnaire, University level</td>
<td></td>
</tr>
<tr>
<td>Ratnadewi and Yunianti (2019) (n = 4)</td>
<td>Indonesia</td>
<td>Muhammadiyah University of Surabaya</td>
<td>Meta-analysis, Observation, University level</td>
<td></td>
</tr>
<tr>
<td>Ow and Tan (2017) (n = 20)</td>
<td>Malaysia</td>
<td>University of Malaya</td>
<td>Experimental, Questionnaire</td>
<td></td>
</tr>
<tr>
<td>Author (Sample Size)</td>
<td>Country</td>
<td>Institution</td>
<td>Design (Measurement Instrument)</td>
<td>Findings</td>
</tr>
<tr>
<td>----------------------</td>
<td>---------</td>
<td>-------------</td>
<td>----------------------------------</td>
<td>----------</td>
</tr>
</tbody>
</table>
| Saputri et al. (2018) (n = 294) | Indonesia | Sebelas Maret University | • Descriptive research  
• Questionnaire  
• Grade 12th, secondary school | they are weak in analyzing, evaluating, applying, and making inference during problem-solving. The critical thinking skill test resulted in the evaluation aspect score that reached the highest score, followed by self-regulation skill and analysis. On the other hand, interpretation skill, inference and explanation skill got the lowest percentage of students’ aspects. |
| Malik et al. (2018) (n = 60) | Indonesia | Indonesia University of Education | • Quasi experimental  
• Questionnaire  
• University level | Improved critical thinking skills in groups that apply the verification labs of three moderately categorized aspects (explain, analyze, and evaluate), while the other three aspects are categorized as low (interpreting, self-regulation, and inference). It can be concluded that junior high school students are still in a position that their critical thinking skills are in enough category. |
| Sarasvati and Sriyati (2018) (n = 40) | Indonesia | Indonesia University of Education | • Experimental  
• Questionnaire  
• Grade 8th, secondary school | |
<table>
<thead>
<tr>
<th>Author (Sample Size)</th>
<th>Country</th>
<th>Institution</th>
<th>Design (Measurement Instrument)</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Setiawan, Malik,</td>
<td>Indonesia</td>
<td>Indonesia University of Education</td>
<td>Quasi experimental</td>
<td>Critical thinking skills aspect classified into two categories, namely medium category for explain, self-regulation and analyze and low category for interpret, inference, and evaluate.</td>
</tr>
<tr>
<td>Suhandi, and Permatasari (2017) (n = 60)</td>
<td></td>
<td></td>
<td>Questionnaire</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>University level</td>
<td></td>
</tr>
<tr>
<td>Hunaidah, Wasis,</td>
<td>Indonesia</td>
<td>State University of Surabaya</td>
<td>Experimental</td>
<td>Positive results indicate an increase in collaborative critical thinking skills of physics education students, which is shown that all indicators of collaborative critical thinking skills are in high category.</td>
</tr>
<tr>
<td>Prahani, and</td>
<td></td>
<td></td>
<td>Questionnaire</td>
<td></td>
</tr>
<tr>
<td>Mahdiannur (2018)</td>
<td></td>
<td></td>
<td>University level</td>
<td></td>
</tr>
<tr>
<td>(n = 56)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Irwanto et al.</td>
<td>Indonesia</td>
<td>State University of Yogyakarta</td>
<td>Quasi experimental</td>
<td>Inference and Analysis skills got the lowest mean score comparing to the other sub skills in critical thinking skills.</td>
</tr>
<tr>
<td>(2018) (n = 48)</td>
<td></td>
<td></td>
<td>Questionnaire</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>University level</td>
<td></td>
</tr>
<tr>
<td>Yulianti et al.</td>
<td>Indonesia</td>
<td>State University of Malang</td>
<td>Experimental</td>
<td>The high average scores are in interpreting, self-regulation, and explanation skill. Inference and Analysis skill got the lowest average score.</td>
</tr>
<tr>
<td>(2018) (n = 25)</td>
<td></td>
<td></td>
<td>Questionnaire</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Grade 11th, Secondary school</td>
<td></td>
</tr>
<tr>
<td>Smith et al. (2019)</td>
<td>USA</td>
<td>Wingate University School of Pharmacy</td>
<td>Quasi experimental</td>
<td>The sub-scores where the students scored highest on the test were explanation and analysis.</td>
</tr>
<tr>
<td>(n = 88)</td>
<td></td>
<td></td>
<td>Questionnaire</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>University level</td>
<td></td>
</tr>
<tr>
<td>Author (Sample Size)</td>
<td>Country</td>
<td>Institution</td>
<td>Design (Measurement Instrument)</td>
<td>Findings</td>
</tr>
<tr>
<td>---------------------</td>
<td>---------</td>
<td>-------------</td>
<td>----------------------------------</td>
<td>----------</td>
</tr>
</tbody>
</table>
| Hussein et al. (2019) (n = 127) | Malaysia | University of Malaya | - Quasi experimental  
- Questionnaire  
- Grade 5th, primary school | Meanwhile, interpretation, inference and evaluation were the lowest in the test result. Explanation skill emerged as the highest mean score, and evaluation skill emerged as the lowest mean score. |
| Ramandha, Andayani, and Hadisaputra (2018) (n = 75) | Indonesia | University of Mataram | - Quasi experimental  
- Questionnaire  
- Grade 10th, secondary school | Interpretation, Analysis, and Evaluation skill show the significance of the criteria for critical thinking skills. Inference and explanation skill have lower critically. |
| Amalia, Hartono, and Indaryanti (2019) (n = 30) | Indonesia | Sriwijaya University | - Descriptive research  
- Trigonometric questions  
- Grade 10th, secondary school | The highest average value is in the indicators of interpretation which has an excellent category. The lowest average value is in the indicator of inference which has the poor category. |
| Fernandi et al. (2018) (n = 110) | Indonesia | Indonesia University of Education | - Descriptive research  
- Questionnaire  
- Grade 9th, secondary school | Analysis skill had the highest mean score compared to interpreting and inference skill. |
| Saprudin, Liliasari, Prihatmanto, and Setiawan (2018) (n = 46) | Indonesia | Indonesia University of Education | - Survey research  
- Questionnaire  
- University level | Analysis skill is in the high category, meanwhile evaluation and |
<table>
<thead>
<tr>
<th>Author (Sample Size)</th>
<th>Country</th>
<th>Institution</th>
<th>Design (Measurement Instrument)</th>
<th>Findings</th>
</tr>
</thead>
</table>
| Cahyana, Fitriani, Rianti, and Fauziyah (2018)            | Indonesia   | State University of Jakarta           | • Qualitative research  
• Questionnaire  
• Grade 10th, secondary school                                    | explanation are in the low category. This good category shows that all students are able to meet the critical thinking ability indicator being studied by the researcher. |
| Yerimadesi et al. (2018) (n = 67)                         | Indonesia   | State University of Padang            | • Quasi-experimental  
• Questionnaire  
• Grade 11th, secondary school                                    | All critical thinking indicators (analysis, inference, and explanation skills) are in the “very good” category. The result shows that the percentage achievement of students’ critical thinking skills in solving the static fluid problem for indicators of assessment, inference, and strategy is still low, and the indicator of clarification is quite high. |
| Pamungkas, Aminah, and Nurosyid (2019) (n = 99)           | Indonesia   | Sebelas Maret University              | • Descriptive research  
• Questionnaire  
• Grade 11th, secondary school                                    | The result showed that assessment, analysis, and inference skills did not significantly increase during the time. The six critical thinking sub-skills identified, only the interpretation sub-skill was in the |
| Shirazi and Heidari (2019) (n = 499)                      | Iran        | Shiraz University of Medical Sciences | • Experimental  
• Questionnaire  
• University level                                                   |                                                                 |
Besides, CTS have six subskills that have been included in this systematic review—interpreting data, analysis, inference, evaluation, explanation, and self-regulation. After conducting data extraction, three subskills were included in a crucial category in science education; the percentages are provided in Figure 4. Interpreting data (55.0% of 20 studies) was found to be one of three crucial skills (Asefi & Imani, 2018; Kumar, 2017; Malik et al., 2018; Saputri et al., 2018; Stephenson et al., 2019).

The second skill was inference with 77.8% of 27 studies revealing a low mean score (Dilekli, 2017; Fernandi et al., 2018; Irwanto et al., 2018; Ow & Tan, 2017; Putra & Prayitno, 2018; Ramos et al., 2013; Shirazi & Heidari, 2019; Smith et al., 2019; Zhou et al., 2013). Most of these studies compared the average score of each subskill after conducting different methods such as inquiry-based learning, formal logic course, peer-lead team learning, problem-based learning, knowledge learning in the digital classroom.

**Crucial Domains in Reasoning Skills**

The last cognitive skill is reasoning with six subskills—

<table>
<thead>
<tr>
<th>Author (Sample Size)</th>
<th>Country</th>
<th>Institution</th>
<th>Design (Measurement Instrument)</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>school</td>
<td></td>
<td></td>
<td>fair category, while the remaining five sub-skills of students’ critical thinking were in a low category. The evaluation, self-regulation, and inference sub-skills were the sub-skills with the lowest percentage. It can be concluded that many students were less capable in those critical thinking sub-skills.</td>
<td></td>
</tr>
</tbody>
</table>
conservation law, proportional, controlling of variable, probabilistic, correlational, and hypothetical deductive reasoning. These skills were measured in 12 studies (three paper proceedings and nine journal articles) with 31,028 total sample number, shown in Table 3.

Different from the two previous skills, all participants in this domain came from two levels only—secondary and university—and all studies were conducted quantitatively.

As shown in Figure 5, the percentage of included studies suggest that all subskills in reasoning were in the poor category, evidenced in none of the skills obtaining more than 50% in the good category. Conservation weight and volume were found in 60% of 10 studies and 66.7% of nine, respectively, thus put under crucial. This was followed by proportional reasoning with 66.7% of nine studies, controlling of variable with 50.0% of 10 studies, probabilistic with 58.3% of 12 studies, correlational with 75% of 12 studies, and hypothetical deductive reasoning skill with 66.7% of 9 studies, thereby categorized as crucial skills in science education (Jensen et al., 2017; Piraksa et al., 2014; Remigio et al., 2014; Ross & Cousins, 2006; Susilawati & Anam, 2017; Wulandari & Shofiyah, 2018; Yuksel, 2019).

<table>
<thead>
<tr>
<th>Author (Sample Size)</th>
<th>Country</th>
<th>Institution</th>
<th>Design (Measurement Instrument)</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Remigio et al., 2014) (n = 93)</td>
<td>Philippines</td>
<td>Ateneo de Manila University</td>
<td>Quasi-experimental • Questionnaire • Grade 10th, secondary level</td>
<td>Conservation of weight and volume, and probabilistic reasoning skill got the highest mean score.</td>
</tr>
<tr>
<td>(Muslim Indo et al., nesia 2017) (n = 104)</td>
<td>Indonesia</td>
<td>Indonesia University of Education</td>
<td>Research and Design • Questionnaire • Secondary school</td>
<td>Proportional reasoning, Control of variable, and Correlational reasoning skill got the lowest mean score. Only hypothetical deductive reasoning got the highest average score.</td>
</tr>
<tr>
<td>(Mendoza et al., 2018) (n = 35)</td>
<td>Columbia</td>
<td>Manuela Beltran University</td>
<td>Quasi-experimental • Questionnaire • University</td>
<td>Conservation, control of variable, probabilistic reasoning, and correlational reasoning got the lowest average score. In PBL, proportional variable skill, control of variable, and probabilistic reasoning skill had shown improvement of correct answer percentage.</td>
</tr>
<tr>
<td>Author (Sample Size)</td>
<td>Country</td>
<td>Institution</td>
<td>Design (Measurement Instrument)</td>
<td>Findings</td>
</tr>
<tr>
<td>----------------------</td>
<td>---------</td>
<td>-------------</td>
<td>----------------------------------</td>
<td>----------</td>
</tr>
<tr>
<td>(Mendoza et al., 2018) (n = 35)</td>
<td>Columbia</td>
<td>Manuela Beltran University</td>
<td>• Quasi-experimental • Questionnaire • University level</td>
<td>Conservation skill, correlational reasoning, and hypothetical-deductive reasoning had decreased of correct answer percentage. In CL, all seven reasoning skills had decreased of correct answer percentage.</td>
</tr>
<tr>
<td>(Piraksa et al., 2014) (n = 400)</td>
<td>Thailand</td>
<td>Khon Kaen University</td>
<td>• Experimental • Questionnaire • Grade 11th, secondary school</td>
<td>Conservation mass and volume skill and correlational reasoning skill showed the highest mean score. Proportional reasoning skills, control of variables, probabilistic reasoning, and hypothetical-deductive reasoning skill show the lowest mean score.</td>
</tr>
<tr>
<td>(Yuksel, 2019) (n = 31)</td>
<td>Turkey</td>
<td>Gazi University</td>
<td>• Experimental • Questionnaire • University level</td>
<td>Conservation laws, correlational reasoning, and hypothetical deductive reasoning got the lowest mean score. Conservation mass and volume, probabilistic reasoning, and control of variables skill got the highest mean percentage score.</td>
</tr>
<tr>
<td>(Stammen et al., 2018) (n = 32)</td>
<td>USA</td>
<td>The Ohio State University</td>
<td>• Experimental • Questionnaire • University level</td>
<td>Proportional reasoning skills, correlational reasoning, and hypothetical deductive reasoning got the lowest mean percentage score. The impact of the program was mediated by teacher commitment to improving students’ correlational reasoning skills and by teacher efficacy. The program was less successful in developing students’ ability to conclude correlational data.</td>
</tr>
<tr>
<td>(Ross &amp; Cousins, 2006) (n = 12)</td>
<td>Canada</td>
<td>Ontario Institute for Studies in Edu</td>
<td>• Experimental • Questionnaire • Grade 9th to 10th, secondary School</td>
<td>Conservation mass and volume, probabilistic reasoning, and control of variables skill got the highest mean percentage score. Proportional reasoning skills, correlational reasoning, and hypothetical deductive reasoning got the lowest mean percentage score. The impact of the program was mediated by teacher commitment to improving students’ correlational reasoning skills and by teacher efficacy. The program was less successful in developing students’ ability to conclude correlational data.</td>
</tr>
<tr>
<td>Author (Sample Size)</td>
<td>Country</td>
<td>Institution</td>
<td>Design (Measurement Instrument)</td>
<td>Findings</td>
</tr>
<tr>
<td>----------------------</td>
<td>---------</td>
<td>-------------</td>
<td>---------------------------------</td>
<td>----------</td>
</tr>
<tr>
<td>(Jensen et al., 2017)</td>
<td>USA</td>
<td>Brigham Young University</td>
<td>Experimental • Questionnaire • University level</td>
<td>Conservation of mass and probabilistic reasoning skill got the highest score of the total mean score. Proportional reasoning, Control of variable, Correlational reasoning, and Hypothetical deductive reasoning got the lowest score of the total mean score.</td>
</tr>
<tr>
<td>(Wulandari &amp; Shofiyah, 2018)</td>
<td>Indonesia</td>
<td>University of Muhammadiyah Sidoarjo</td>
<td>Experimental • Questionnaire • University level</td>
<td>Proportional reasoning and control of variables got the highest student’s mastery. Conservation laws, Probabilistic reasoning, and correlational reasoning got the lowest student’s mastery.</td>
</tr>
<tr>
<td>(Susilawati &amp; Anam, 2017)</td>
<td>Indonesia</td>
<td>State University of Semarang</td>
<td>Pre experimental • Questionnaire • Grade 11th, Secondary school</td>
<td>Hypothetical deductive reasoning got the highest mean score with high increment. Correlational reasoning and Probabilistic reasoning skill got the lowest mean score.</td>
</tr>
<tr>
<td>(Rosdiana et al., 2019)</td>
<td>Indonesia</td>
<td>Indonesia University of Education</td>
<td>Observation • Questionnaire • Grade 9th, Secondary school</td>
<td>Combinatorial reasoning, Correlational reasoning, and Controlling variables got a high percentage of correct answer. Conservation reasoning, proportional reasoning, and probabilistic reasoning got the lowest percentage of the correct answer.</td>
</tr>
</tbody>
</table>

Figure 4. Summary of the findings of 34 included studies in CTS.
In a total, eight studies revealed the range of mean scores in order to identify the category of subskills with and without intervention. Six of eight studies were conducted to analyze learning methods such as inquiry-based learning, problem-based learning, 5E learning model, analogy-enhanced instruction. Extrapolate the findings of RS in science education.

**The Crucial Skills among SPS, CTS, and RS**

Our literature search revealed some articles which showed the relationships among cognitive skills in science education. It was found that the most crucial subskills are in groups 2, 3, and 4, as shown in Figure 6, starting from group 2 that pictures relationships between SPS and RS. In this group, it was shown that the RS covered most of the subskills in integrated SPS, including identification and controlling variable (HD2), formulating hypothesis.
The findings from this systematic review showed that these skills are crucial skills between SPS and RS. Second, group 3 demonstrates the relationship among SPS, RS, and CTS including interpreting data (HD1), inference (HD4), and analysis skills (HD4). The last group demonstrates the relationship between RS and CTS consisting of simple explanation (conservation reasoning/EI), explanation (HD5), and evaluation (HD5). These skills are covered under RS and correlation in order to improve students’ achievements. The previous study revealed that learning processes should contain some learning activities that challenge students’ CTS, RS and improve their SPS in problem-solving, finding, and analyzing to establish an appropriate concept (Naimnule & Corebima, 2018).

CONCLUSION
In conclusion, the issue of cognitive skills raises important questions regarding the crucial skills for the cognition stage—SPS, CTS, and RS. SPS focuses on the whole learning process including basic and integrated SPS. Based on the findings in this systematic review, the crucial subskills are inference, measuring, identifying and controlling variable, definition operational variable, and explanation, which mostly consisted of the integrated domain. Also, CTS focuses on the evaluation of the learning process with crucial subskills, including interpreting data, inference, and evaluation. Furthermore, RS is an essential element in the learning process found under the crucial domain. In this domain, students start giving explanations or reasons with logical thinking for each hypothesis, statement, data, opinions, theory, experimental design, conclusion, etc. Moreover, after identifying their relationships, we found that most of the crucial subskills among these three cognitive skills were mostly under RS, covering process and critical thinking skills in one circle, shown in Figure 6. It is connected to three points: (i) Most of the subskills in the integrated science process skill, (ii) interpreting data skill (HD1), analysis skill (HD4) and inference skill (HD4), (iii) explanation skill (EI/HD5) and evaluation skill (HD5). Based on this finding, it can be concluded that crucial cognitive skills in science education are in the RS domain.

REFERENCES


Han, J 2013, ‘Scientific Reasoning: Research, Development, and Assessment’, Doctor of Philosophy Dissertation, The Ohio State University, Ohio.


Hasanah & Shimizu
through Process-Oriented Guided-Inquiry Learning (POGIL’), *International Journal of Instruction*, vol. 11, no. 4, pp. 777-94.


Kumar, RR 2017, ‘Effectiveness of Formal Logic Course on the Reasoning Skills of Students in Nizwa College of Technology, Oman’, *Journal of Education and Practice*, vol. 8, no. 7, pp. 30-5.


Osman, K & Vebrianto, R 2013, ‘Fostering Science Process Skills and Improving Achievement Hasanah & Shimizu'
through the Use of Multiple Media’, *Journal of Baltic Science Education*, vol. 12, no. 2, pp. 191-204.


Shin, KR 1998, ‘Critical Thinking ability and Clinical Decision-making Skills among Senior Nursing Students in Associate and Baccalaureate Programs in

Hasanah & Shimizu


