Higher Order Thinking and Critical Thinking Skills in Problem-Based Learning Environments: A Systematic Review

Rahma Hidayati Darwis1 ✉ Alimuddin2 and Andi Patimbangi3
1WIIN Bone, South Sulawesi, Indonesia
2UNM, South Sulawesi, Indonesia
Corresponding Author: Rahma Hidayati Darwis, E-mail: rahma_darwis@yahoo.com

ABSTRACT
Critical thinking is widely considered an important competency to acquire in education. Learner exposure to problems and collaboration has been shown to help in improving critical thinking processes. These elements are contained in a learner-centered learning environment, such as problem-based learning. In addition to critical thinking, higher order thinking skills also contain elements of PBL. However, the definitions of HOT and critical thinking are often poorly defined and overlapping. Therefore, this SJR research aims to investigate how HOTs and critical thinking are conceptualized in a PBL environment. Another objective of this study was to review the evidence of the effectiveness of the PBL environment in developing HOTs and critical thinking. The results showed that from various references, it was found that the focus of the study was more on the process of critical thinking rather than critical thinking disposition (i.e., the tendency or willingness to engage in critical thinking). Furthermore, despite the positive effects of PBL on HOTs and critical thinking, limitations in the findings include the lack of clarity and consistency in how researchers conceptualize and measure these forms of thinking, both HOTs and critical thinking. In addition, important components of PBL are often overlooked. Finally, the study also identified various design issues in the effects study, such as the lack of a control group, which makes the results of the investigation questionable.

KEYWORDS
PBL, higher order thinking, critical thinking.

ARTICLE INFORMATION
ACCEPTED: 02 May 2024  PUBLISHED: 30 May 2024  DOI: 10.32996/jlds.2024.2.2.3

1. Introduction
Higher order thinking is generally considered a very important skill for learners to learn. The urgency of such thinking has increased over time (Pellegrino & Hilton, 2012). Mastery of higher-order thinking skills is no longer limited to learners but also to working professionals and the entire knowledge society (Bezanilla et al., 2021). Therefore, higher-order thinking skills are one of the goals of education today (e.g., Butler & Halpern, 2020). A metaanalysis showed that opportunities for dialogue, exposing learners to authentic problems and examples, and guiding learners through habitual thinking had a significant impact on learners’ higher-order thinking skills. (Abrami et al., 2015). These activities are in a learner-centered learning environment. Therefore, it seems that the best way to improve learners’ higher order thinking skills is to use active and learner-centered learning methods (Bezanilla et al., 2021; Lombardi et al., 2021). Problem-based learning (PBL) is one of the prototypical examples of active, learner-centered learning methods. In addition to higher-order thinking (HOT), it contains elements that are present in learner-centered learning environments, such as PBL.

For the aforementioned reasons, in this study, the focus is on how to place higher order thinking skills in problem-based learning, considering that other terms such as critical thinking and analytical thinking also contain elements that exist in a learner-centered learning environment and are often unclear and overlapping. This research focuses on problem-based learning as the PBL learning
format is most frequently studied in the research literature, and it is also one of the learner-centered learning (Author, 2022; Nagarajan & Overton, 2019). In addition, PBL is one of the learning models that has been included as a recognized learning format in the Cambridge Handbook of Learning Sciences (Sawyer, 2014) and has specific criteria that must be met to be labeled as PBL. Despite its unique characteristics and origins, PBL as a pedagogy can be seen as an embodiment of constructivist learning, which is one of the current learning paradigms. Constructivism is a theory or view of how learning occurs, which states that learners construct knowledge from experience. It is rooted in philosophy and emphasizes the active role of learners in their knowledge acquisition process (Loyens et al., 2012).

1.1 What is PBL?
There are many definitions of PBL in the research literature, all presenting different perspectives and ideas with respect to educational pedagogy (Dolmans et al., 2016; Zabit, 2010). Despite the variety of definitions, problem-based learning exhibits some common characteristics in its implementation. Based on the original method developed at McMaster University (Spaulding, 1969), Barrows (1996) explained that there are at least 6 (six) core characteristics of problem-based learning, namely: first, learning is learner-centered; second, learning occurs in small groups of learners under the guidance of a tutor; third tutor or teacher as a facilitator or guide; fourth, learners face so-called contextual problems (i.e., related to real life); fifth, these problems serve as a stimulus for activating learners' prior knowledge which leads to the discovery of knowledge gaps. Overcoming this knowledge gap is done through self-directed learning, which requires sufficient time for self-learning (Schmidt et al., 2009).

In addition to the core elements, Wijnia et al. (2019) highlight that PBL as a process consists of three separate stages: the initial discussion stage, the self-study stage, and the reporting stage. First, learners are presented with a meaningful problem that describes an observable phenomenon or event. The learning objectives of the problem presented to learners can vary. For example, the problem could come from professional practice or be directly related to a specific event in a particular domain or field of study. An example of a problem related to a specific domain of study in field X is as follows: “Coming home from work, Anita's body feels tired and needs a hot bath; Anita, a manager in a company, finds two spiders in her bathtub. She shrunk back, screamed, and ran away, her heart pounding, cold sweat pouring all over her body. A neighbor saved her from her predicament by killing the small animals with a newspaper.” What does this tell us about the incident?

In the first stage, prior knowledge plays an important role, as learners generate theories to explain the problem based on their life experiences. As their knowledge is often limited and insufficient, learners formulate a learning problem (formulated as a question) to guide their study and further self-learning. The activity takes the form of a class discussion, usually in classes with fewer than 12 learners. In the second stage, learners consult learning resources to gain knowledge relevant to the problem and to answer the learning problem questions. Learning resources can be chosen by learners, tutors, or a combination of both (Wijnia et al., 2019). Alongside these steps, learners should plan and monitor learning activities that need to be done before the next class meeting (Loyens et al., 2008). In the final stage, learners regroup under the guidance of a tutor or teacher to share and critically evaluate their findings and elaborate on their newly acquired knowledge. Learners apply this knowledge to the problem to identify a plausible solution or explanation (Loyens & Rikers, 2017; Wijnia et al., 2019).

1.2 What PBL fosters
One of the purposes of schools and colleges implementing learner-centered approaches such as PBL is to improve learners' competence in addressing complex problems that are common in a changing world (Gijbels et al., 2005). To that end, several goals and desired outcomes have been put forward for PBL. The main goal is to educate learners to a level where they can comfortably use and retrieve information when needed and identify situations where specialized knowledge and strategic processes can be applied. With these strategies and knowledge, learners can begin to develop plausible explanations of phenomena that represent important disciplinary understanding (Loyens et al., 2012; McNeill & Krajcik, 2011). Several studies have investigated the effectiveness of PBL on knowledge acquisition (e.g., Chen & Yang, 2019; Strobel & Van Barneveld, 2009). In addition, PBL consists of collaborative learning sessions that can foster effective interpersonal communication. Such skills can enable learners to contribute to discussions in a clear and precise manner, help reach conclusions and answers more easily, and identify inconsistencies and unresolved issues (Loyens et al., 2008, 2012).

Furthermore, learners can develop problem-solving strategies while working on the problem or project (Krajcik et al., 2008). Although the problems are highly complex and unstructured, the responses can be effectively analyzed, and reasonable responses can be identified (Loyens et al., 2012). In addition, since problems and projects are specific to learners' field of study, the knowledge and strategies they acquire can be applied to their future professional practice. Therefore, problems and projects are believed to be more interesting, motivating and engaging for learners (Hmelo-Silver, 2004; Larmer et al., 2015; Saad & Zainudin, 2022). Finally, as already mentioned, learner-centered teaching methods such as PBL imply a different and less directive role for the teacher. As a result, learners receive a lot of knowledge during the learning process. Therefore, the success of problem-based learning is also largely determined by "learners' readiness to engage in learning activities that are determined by themselves, not by the teacher".
Although the research literature on PBL does not explicitly state that such learning should encourage higher-order thinking, its design and implementation seem to require the involvement of learners in such forms of thinking. Thus, in this review, the researcher sought to define the construct in the context of a learner-centered learning environment.

1.3 What is higher order thinking?

*Higher order thinking* (HOT) can be seen as an overarching concept defined as "skills that enhance the construction of deeper, conceptually driven understanding" (Schraw & Robinson, 2011, p. 2). Framed in more conventional terms, higher order thinking (HOT) corresponds to Bloom’s taxonomy, with a recall of facts reflecting cognitive thinking (i.e., relating to the acquisition of knowledge or information), understanding, applying, analyzing, synthesizing, and evaluating as higher order thinking, referring to more intellectual abilities and skills (Lombardi, 2022; Miri et al., 2007). The focus on thinking skills does not mean that the importance of knowledge is neglected. Knowledge is indispensable and related to the thinking process (Lombardi, 2022).

*High order thinking* (HOT) in various references has four main components (Schraw et al., 2011): 1) reasoning (i.e., induction and deduction), 2) argumentation (i.e., producing and evaluating evidence and arguments), 3) metacognition (i.e., thinking about and organizing one’s thinking), and 4) problem solving and critical thinking. Problem solving involves several steps that are performed sequentially: 1) identifying and representing the problem at hand, 2) selecting and applying appropriate solution strategies, and 3) evaluating the process and solution (Chakravorty et al., 2008). Meanwhile, critical thinking refers to reflective thinking that leads to specific outcomes (i.e., decision-making) and actions (Ennis, 1987). Critical thinking, in that view, is considered a subcomponent of higher order thinking (HOT) (Schraw et al., 2011). Indeed, in its measurement, critical thinking has also addressed HOT processes such as analysis and synthesis (Lombardi, 2022). Yen and Halili (2015) also characterized HOT as an umbrella term for all ways of reflective thinking, including critical thinking, problem solving, decision making, and metacognitive processing. In addition, in this definition, some components refer to Schraw and colleagues’ (2011) taxonomy: reflection (component 4) and metacognition (component 3), problem solving (component 4), and decision making (critical thinking, component 4). The only exception is creative thinking, which is not included as a subcomponent by Schraw and colleagues (2011); however, they acknowledge that it could be part of a broader HOT taxonomy.

1.4 What is critical thinking?

Tracing back to the Greek philosophers who sought to explain the origin and meaning of critical thinking. As Van Peppen (2020) points out, "the word critical comes from the Greek words 'kritikos' (i.e., to judge/discard) and 'kriterion' (i.e., standard)," and hence "critical thinking means making judgments based on standards." The second thinker on critical thinking is John Dewey, who talks about "reflective thinking" when referring to critical thinking. The second thinker on critical thinking was John Dewey, who spoke of "reflective thinking" when referring to critical thinking. From there, many traditions and definitions of critical thinking were formulated. Ennis (1987), for example, defines critical thinking as a thinking process that focuses on decisions about what to believe or do. The view was expanded through the ideas of Glaser (1941), who recognized the role of disposition in critical thinking. Ennis (1962) distinguishes two different components of critical thinking, namely dispositions and abilities, with the former being more trait-like tendencies (e.g., disposition toward curiosity, open-mindedness, sensitivity to other points of view, cognitive flexibility) and the latter, abilities, referring to actual cognitive activities (e.g., focusing, analyzing arguments, asking questions, evaluating evidence, comparing potential outcomes; Schraw et al., 2011).

Various scholars from the American Philosophical Association tried to come up with a consensus regarding the definition of critical thinking. As explained by the Delphi Panel (Facione, 1990b), the processes associated with critical thinking are interpretation (i.e., understanding and articulating meaning), analysis (i.e., identifying relationships between information, including analysis of arguments), evaluation (i.e., making judgments and assessments about the credibility of information, including assessing arguments), inference (i.e., identifying information necessary for decision making, including generating hypotheses), explanation (i.e., articulating and presenting one’s position, arguments, and the analysis used to determine that position), and self-regulation (i.e., self-analyzing and checking one’s conclusions and correcting when necessary; Facione, 1990b). The last component (i.e., self-regulation) has a strong meta-cognitive character (Zimmerman & Moylan, 2009).

Another conceptualization of critical thinking that resulted in the development of a widely used critical thinking measurement instrument came from Halpern. Halpern (2014) focused more on cognitive abilities/activities and less on dispositions in defining critical thinking. Specifically, Halpern states that critical thinking entails "cognitive skills or strategies that increase the likelihood of desired outcomes". Halpern’s taxonomy consists of five main elements: verbal reasoning, argument analysis, hypothesis testing, equations and uncertainty, and decision making/problem solving.

1.5 HOT and critical thinking links

To establish the relationship between HOT, critical thinking and PBL learning environments, the researcher looked at two lines of research literature: literature on how to effectively teach HOT critical thinking, as well as literature on the learning processes involved in PBL, such as the synthesis between cognitive science and learning). Much of the research directed at teaching forms of
Higher Order Thinking and Critical Thinking Skills in Problem-Based Learning Environments: A Systematic Review

Reflective thinking has addressed critical thinking as a form of HOT (e.g., Abrami et al., 2015; Miri et al., 2007; Schraw et al., 2011). For example, Miri and colleagues (2007) defined three teaching strategies that should encourage learners to engage collaboratively in processes aligned with critical thinking (e.g., asking the right questions and seeking reasonable solutions). These teaching strategies are (a) addressing real cases in class, (b) encouraging class discussions, and (c) encouraging inquiry-oriented experiments.

The link with PBL is obvious, as PBL centers on addressing real problems and cases in collaborative classroom discussions. Of course, PBL sometimes requires experiments to be conducted. In a meta-analysis, Abrami and colleagues (2015) reviewed possible learning strategies that can promote critical thinking, concluding that there are two types of interventions that help develop critical thinking processes: discussion and "authentic or situational problems or examples, especially when problem solving is applied or used" (Abrami et al., 2015, p. 302). Clearly, there is a clear connection as the use of authentic problems and class discussions are key components of PBL. In addition, Torf (2011) calls the core activities of PBL "high critical thinking activities".

Other research also reveals the relationship between HOT, critical thinking and PBL in the learning process. For example, Hung, W. and colleagues (2008) showed that learners who experienced problem-based learning had better hypothesis testing ability due to their more coherent hypothesis explanations and hypothesis-driven reasoning. Furthermore, the PBL process relies heavily on group discussion of real-life problems, discovering knowledge gaps, gathering information/evidence to answer the learning issue/question, analyzing evidence, resolving ambiguities, and deciding on the outcome.

1.6 The Present Study
Although the development of HOT and critical thinking may not be an explicit goal of a learner-centered approach, there are theoretical and empirical reasons that suggest a correlation between HOT and critical thinking and problem-based learning. Literature relating to research on how to effectively teach HOT and critical thinking has suggested that both skills can be enhanced when learning is formulated using discussion and problem solving. Other research literature also discusses learning processes that take place in learner-centered learning environments such as PBL, mentioning that HOT and critical thinking processes are manifested in the form of decision making, argumentation, weighing evidence, explanation, investigation (Abrami et al., 2015; Krajcik et al., 2008). Therefore, the first research question that the researcher posed was: how are HOT and critical thinking conceptualized in a learner-centered learning environment? Moreover, there are similarities between the processes involved in HOT, critical thinking on the one hand and the learning process in PBL on the other. In addition, effective learning activities to foster HOT and critical thinking are core activities in PBL. Therefore, the second objective of this study was to review the evidence of the effectiveness of learner-centered environments in developing HOT and critical thinking. To that end, the researcher conducted a review of studies investigating HOT critical thinking in the context of PBL.

2. Methodology
2.1 Search strategy
For this review, six online databases were systematically investigated, namely Web of Science (Core Collection) and five EBSCO databases (ERIC, Medline, PsycInfo, Psychology and Behavioral Sciences, and Teacher Reference Center). In this study, Medline was included in the list because problem-based learning originated in medical education (Spauld-ing, 1969; see also ServantMiklos, 2019) and is often researched in the context of medical education (W. Hung et al., 2019; Kohl et al., 2019). Hung et al., 2019; Koh et al., 2008; Smits et al., 2002; Strobel & Van Barneveld, 2009). For this systematic review, the researcher used the following set of Boolean search terms (Oliver, 2012): “problem-based learning”, or “problem-based approach”, or “problem-based instruction”, or “PBL”, and “higher-order thinking", or "critical thinking". Due to its medical connotations (i.e., it also stands for "peripheral blood lymphocytes"; e.g., Caldwell et al., 1998), "PBL" was not used for the Web of Science search. The search term fits this research question as it encompasses PBL as well as HOT, and critical thinking, thus leading the researcher to studies incorporating these variables. In addition to using the terms mentioned above, the researcher limited the search to peer-reviewed notes written in English.

2.2 Selection process
2.2.1 Inclusion criteria
In addition to the search parameters set by the search terms in this study, inclusion criteria terms were also used to determine the final sample of the study with the following considerations: 1) using quantitative measures of HOT critical thinking, 2) conducted in different locations in a PBL environment, 3) being an empirical study that takes place in a classroom context, 4) investigating learners, and 3) published as a peer-reviewed journal article. For example, to be included in our review, studies must have specific measures of HOT and critical thinking that provide insight into the researcher’s conceptualization of the constructs. The measure used can be a standardized instrument or one designed by the researcher. In addition, in line with the research objectives, we only included studies that focused on PBL and met the basic criteria for such a learner-centered approach. To assess the quality of these learning environments, the definitions or descriptions in the theoretical framework and the implementation of these environments
were assessed based on defining characteristics set out in the literature (Barrows, 1996; Hmelo-Silver, 2004; Schmidt et al., 2009). The criteria include 1) learner-centered active learning, 2) the role of the teacher as a guide, 3) collaborative learning in small groups, 4) the use of realistic problems at the beginning of the learning process, and 5) sufficient time for self-directed learning. Furthermore, PBL must contain three phases of the process (i.e. initial discussion, self-study, and reporting phases). In this research, only studies in the context of classroom and higher education are included as they focus on the effectiveness of learning environments in formal education settings.

2.2.2 Exclusion criteria
In finding study materials, the researcher excluded items such as theses or self-reports (often from course evaluations) as well as excluding the critical thinking scale from a Motivation for Learning Strategies Questionnaire (MSLQ; Duncan & McK-eachie, 2005). In the MSLQ, critical thinking is viewed as a learning strategy. Researchers have excluded these studies (e.g., Sungur & Tekkaya, 2006). This research has also excluded studies that did not meet the criteria for PBL. For example, if a study was indicated as investigating PBL where the learning process started with a lecture rather than being based on a contextual problem. In addition, it also excluded studies that did not adequately describe the learning process by the researchers yet were labeled as PBL (e.g., Razali et al., 2017). Also excluded were studies that consisted of combining PBL with additional activities (e.g. concept maps with PBL; Si et al., 2019) or interventions. In addition, exclusions have also been made for studies conducted in laboratories, intervention studies in venues such as tutoring after-school programs, or studies with employee samples (e.g., health nurses; T.-M. Hung et al., 2015), as the focus on these studies is formal education. Finally, for the review of this study, exclusion was also made for peer-reviewed conference papers and abstracts (“wrong format”) as they were often difficult to retrieve or provided too little information to determine outcome measures and learning environments.

2.2.3 Coding and Final Sample
Figure 1 outlines the overall search process based on the PRISMA (Preferred Reporting Items for Systematic Review and Meta-Analysis) guidelines (Moher et al., 2015). As can be seen in the figure, an initial search was conducted on Google Scholar, Web of Science and EBSCO databases, and 2,968 results were obtained. After removing duplicates identified through Turnitin, 2,545 papers remained, for which the title, abstract and, where necessary, full text were also screened. The researchers identified 27 additional duplicates and excluded 2,405 papers as they did not meet the inclusion criteria. Subsequently, 113 studies were selected for further screening and coding. The specific codes were in line with the formulated inclusion and exclusion criteria and were designed to simplify the final elimination stage (i.e., excluding articles after in-depth reading). The full details can be seen in Table 1, which provides a detailed overview of the codes used.

3. Result and Discussion
Before answering this research question, we first describe the characteristics of the selected studies (see Table 2). From the various studies or research, there are 22 research results that have a high enough relevance to be investigated.
Table 1. Coding scale

<table>
<thead>
<tr>
<th>Category</th>
<th>Specific</th>
</tr>
</thead>
<tbody>
<tr>
<td>High level thinking, critical thinking</td>
<td>Labels (HOT, CT)</td>
</tr>
<tr>
<td></td>
<td>Definition, conceptualization</td>
</tr>
<tr>
<td></td>
<td>Reference</td>
</tr>
<tr>
<td>Learning environment</td>
<td>Definition, conceptualization</td>
</tr>
<tr>
<td></td>
<td>Reference</td>
</tr>
<tr>
<td></td>
<td>Labels (PBL)</td>
</tr>
<tr>
<td></td>
<td>Definition, conceptualization</td>
</tr>
<tr>
<td></td>
<td>Application</td>
</tr>
<tr>
<td></td>
<td>Score based on compliance with established quality criteria (1: not included; 2: Possible; 3: included)</td>
</tr>
<tr>
<td>Academic domain</td>
<td>STEAM/mathematics education/others</td>
</tr>
<tr>
<td>Research questions or objectives</td>
<td>The research questions/objectives are reported in the paper.</td>
</tr>
</tbody>
</table>

Figure 1. Flow diagram of the search and selection process
Three studies were conducted in the Mathematics Education domain, 6 in the Science, Technology, Engineering, Arts, and Mathematics (STEAM) domain, and 11 in other domains (e.g., financial management or psychology). The studies included 11 different countries. Most of the studies were conducted in the United States (n=4) and Indonesia (n=5). All studies were related to PBL (see Table 2).

### 3.1 Research Question 1: Conceptualization of HOT and critical thinking

The first research question of this review is, “How are HOT and critical thinking conceptualized in a learner-centered learning environment?” One of the learner-centered learning models is problem-based learning (PBL).

<table>
<thead>
<tr>
<th>Article</th>
<th>treatment</th>
<th>Construct</th>
<th>instrument</th>
<th>Duration</th>
<th>Domain</th>
<th>Country</th>
</tr>
</thead>
<tbody>
<tr>
<td>Carriger (2016)</td>
<td>PBL</td>
<td>CT</td>
<td>Researcher-developed assignment and rubric based on the association of American Colleges and Universities CTDI and CTTST (Portuguese versions)</td>
<td>1 Semester</td>
<td>Other</td>
<td>USA</td>
</tr>
<tr>
<td>da Costa Carboigim et al. (2018)</td>
<td>PBL</td>
<td>CT</td>
<td>Researcher-developed based on Ennis CTDI</td>
<td>3 Meetings</td>
<td>STEAM</td>
<td>Brazil</td>
</tr>
<tr>
<td>Dakabesri and Luoise (2019)</td>
<td>PBL</td>
<td>CT</td>
<td>Critical Thinking Evaluation Package (Ennis)</td>
<td>6 Meetings</td>
<td>STEAM</td>
<td>Indonesia</td>
</tr>
<tr>
<td>Ding (2016)</td>
<td>PBL</td>
<td>CT</td>
<td>CTDI (researcher-adapted)</td>
<td>9 Weeks</td>
<td>Other</td>
<td>China</td>
</tr>
<tr>
<td>Fitriani et al. (2020)</td>
<td>PBL</td>
<td>CT</td>
<td>CTDI</td>
<td>5 Months</td>
<td>STEAM</td>
<td>Indonesia</td>
</tr>
<tr>
<td>Muehlenkamp et al. (2015)</td>
<td>PBL</td>
<td>CT</td>
<td>Researcher-designed scales: Higher thinking, process, tools, and analysis</td>
<td>16 Weeks</td>
<td>Other</td>
<td>USA</td>
</tr>
<tr>
<td>Rahmat and Hartley (2020)</td>
<td>PBL</td>
<td>CT</td>
<td>Test of Critical Thinking</td>
<td>16 Weeks</td>
<td>STEAM</td>
<td>USA</td>
</tr>
<tr>
<td>Saputro et al. (2020)</td>
<td>PBL</td>
<td>CT</td>
<td>critical thinking questionnaire</td>
<td>6 Weeks</td>
<td>STEAM</td>
<td>Indonesia</td>
</tr>
<tr>
<td>Sasson et al. (2018)</td>
<td>PBL</td>
<td>CT &amp; HOT</td>
<td>Author-constructed test based on dori et al. (2003) and author-constructed test e.g. based on Blooms Taxonomy</td>
<td>2 Year Program</td>
<td>Other</td>
<td>Israel</td>
</tr>
<tr>
<td>Siwe and Mapela (2016)</td>
<td>PBL</td>
<td>CT</td>
<td>Researchers Constructed Test of Science Critical Thinking</td>
<td>18 h</td>
<td>STEAM</td>
<td>Malaysia</td>
</tr>
<tr>
<td>Sugeng and Suryani (2020)</td>
<td>PBL</td>
<td>HOT</td>
<td>Author designed test based on Bloom’s Taxonomy</td>
<td>3 Months</td>
<td>Other</td>
<td>Indonesia</td>
</tr>
<tr>
<td>W.-C.W.Yu et al. (2015)</td>
<td>PBL</td>
<td>CT</td>
<td>CTDI</td>
<td>18 Weeks</td>
<td>Other</td>
<td>Taiwan</td>
</tr>
</tbody>
</table>
Based on the results of the source search, HOT was only investigated in four studies. Most of the other studies investigated the effect of PBL on critical thinking skills. Therefore, to answer this research question, the first study focused on the conceptualization of critical thinking, followed by HOT. The first review is a review of how critical thinking and HOT are defined and measured.

### 3.2 Conceptualization of critical thinking

The conceptualization of critical thinking in PBL consists of critical thinking dispositions and critical thinking processes. These processes are expressed as “skills” or “abilities” in the research conducted. The review in this study uses the term "process" instead of "skill". Furthermore, the term critical thinking disposition can be defined as a constant internal motivation to face problems and make decisions using critical thinking” (Facione et al., 2000, p. 65). Facione et al. (2000) further explained that the use of the word disposition aims to refer to the characterological attributes of individuals. Examples of dispositions are open-minded, analytical, or truth-seeking. In the studies conducted, a disposition is described as the "will" or "tendency" to evaluate situations critically (e.g., Temel, 2014; W.-CW Yu et al., 2015) and is a necessary prerequisite for the critical thinking process. (Temel, 2014)

Although the relationship between dispositions and critical thinking processes is not very high or the correlation between the two is less significant (e.g., Facione et al., 2000), both seem to be necessary for "sensible, reflective thinking focused on making decisions about what to believe or do" (e.g., Ennis, 2011, p.10). Studies that reviewed the definition of the concept of critical thinking at least seven times reviewed dispositions and processes; 13 times were limited to process reviews, and 4 (four) times only in terms of dispositions. Three studies did not define the concept or only spoke generally as “a way to find meaning in real life” (Burris & Garton, 2007, p. 106). Of the studies reviewed, 5 (five) studies measured critical thinking disposition, 10 studies measured critical thinking process, and one study measured both disposition and critical thinking process. A study has also explained that in their activities, researchers have measured critical thinking disposition, but in the results section, only statistical figures for critical thinking processes are reported (Hassanpour Dekhordi & Heydarnejad, 2008). Regarding measurement, the 13 studies align in defining critical thinking and measuring critical thinking. This means that, for example, when the authors defined critical thinking in relation to dispositions, they also measured dispositions. This also means that 8 studies were not aligned in this regard (for example, mentioning process but measuring disposition or mentioning both but measuring one component).

In summary, although critical thinking consists of dispositions and processes, in the conceptualization of critical thinking, especially in PBL research, it appears that most studies only focus on processes, and only a small number combine with dispositions. In addition, the studies were not aligned in the focus (disposition and/or process) of the critical thinking description and the focus of the measurement instrument. Critical thinking dispositions in the studies included in this review have many different terms used to describe critical thinking dispositions. Some studies that reported measuring critical thinking disposition used the California Critical Thinking Disposition Inventory (CCTDI; Facione, 1990a). The CCTDI assesses learners’ willingness or tendency to engage in critical thinking. The CCTDI contains seven dispositions (Facione, 1990a; Yeh, 2002). The truth-seeking scale refers to a mindset that is objective and honest, and seeks the truth even if the findings do not support one’s opinions/interests. Open-mindedness...
refers to tolerance, open-mindedness to conflicting views, and sensitivity to possible biases in oneself. The Analyticity subscale concerns the disposition to anticipate possible consequences, outcomes, and problematic situations. The fourth subscale, Systematicity, measures an organized, orderly, and focused approach to problem solving. Confidence in one's reasoning process is measured in the Self-Confidence critical thinking subscale.

Some of the terms used to describe critical thinking processes can be seen from the three most frequently used instruments, many of which appear as specific components of the test. For example, the most commonly used instruments were the California Critical Thinking Skills Test (CCTST; 3 studies) and the Watson-Glaser Critical Thinking Appraisal (WGCTA; 1 study), and one study used the Cornell Critical Thinking Test (CCTT), but the theoretical framework associated with this test was used in three other studies to measure critical thinking processes. The CCTST, WGCTA, and CCTT instruments are well-known commercial standardized critical thinking measures. The CCTST is actually a companion test to the CCTDI and measures five critical thinking processes: Analysis, Evaluation, Inference, Deductive Reasoning, and Inductive Reasoning (Facione, 1991). Analysis refers to accurately identifying issues and processes such as categorization, deciphering meaning, and clarifying meaning. Evaluation relates to the ability to judge the credibility of statements and the strength of arguments. The Inference subscale measures the ability to draw logical and justifiable conclusions based on evidence and reasoning. Deductive Reasoning relies on strict rules and logic, such as determining the consequences of a particular set of rules, conditions, principles or procedures (e.g. syllogism, mathematical induction). Finally, Inductive Reasoning refers to sound judgment in uncertain, risky, or ambiguous contexts.

Another well-validated test used in the studies included in this review is the WGCTA (Watson & Glaser, 1980). The WGCTA provides problems and situations that require CT skills. It measures CT as a composite of inquiry attitudes (i.e., recognizing the existence of a problem and accepting the need for evidence), knowledge (i.e., about valid inferences, abstractions, and generalizations), and skills in applying these attitudes and knowledge (Watson & Glaser, 1980, 1994, 2009).

This scale consists of five subscales: Inference, Recognition of Assumptions, Deduction, Interpretation, and Evaluation of Arguments. Inference subscale measures the extent to which participants can determine the correctness of conclusions drawn from given data. Assumption Recognition is concerned with recognizing presuppositions or assumptions implicit in statements or assertions. The Deduction subscale measures the ability to determine whether inferences need to be drawn from the given information. Interpretation is concerned with weighing the evidence and deciding whether generalizations based on the given data are justified. Finally, the Argument Evaluation subscale measures the ability to distinguish strong and relevant arguments from weak and irrelevant arguments. The long version of the scale consists of 80 items (parallel forms A and B; Watson & Glaser, 1980), and the short version (From S) contains 40 items (Watson & Glaser, 1994). A more recent testing version is available (Watson & Glaser, 2009); however, the included studies rely on the older, abbreviated version (Burris & Garton, 2007; Şendağ & Odabaşı, 2009).

This test measures five latent dimensions: Induction, Deduction, Observation, Credibility, and Assumption (Leach et al., 2020). However, these five dimensions are reduced to four parts in the test manual. Two dimensions are taken together (i.e. observation and credibility), and some items from one dimension are counted as elements of another section (Leach et al., 2020). Bataineh and Zghoul (2006) described the CCTT level Z subscale in more detail. According to them, CCTT level Z measures six dimensions: Deduction, Semantics, Credibility of Induction, Identification of Definitions and Assumptions, and Identification of Assumptions. The Deduction subscale measures the extent to which one can detect valid thinking. Its subscale, Semantics, measures the ability to access verbal and linguistic aspects of arguments. Credibility concerns the extent to which participants can estimate the truth of a statement. The subscale induction refers to the ability to judge the best inferences and predictions. Definition and Assumption Identification measures the extent to which one can identify the best definition of a given situation. Finally, assumption identification asks participants to select the most likely assumption not mentioned in the text.

3.3 HOT Process

In several studies that measured CT, the authors mentioned that CT is a component of HOT (Cortázár et al., 2021; Dakabesi & Louise, 2019; Sasson et al., 2018). Only four out of 20 studies measured HOT (Sasson et al., 2018; Sugeng & Suryani, 2020). In all four cases, the authors only defined HOT with reference to the more complex levels of thinking in Bloom et al.’s (1956; Krathwohl, 2002) taxonomy for cognitive domains: application, analysis, synthesis and evaluation. In Sasson et al. (2018), comprehension is also included as one of the higher cognitive processes. In contrast, in Sugeng and Suryani’s (2020) research, comprehension was treated as a form of lower-level thinking. All four studies used Bloom’s Taxonomy as a framework for coding learners’ work. Therefore, we mainly look at the same term to conceptualize HOT and also the correspondence between the characterization and measurement of HOT processes based on Bloom’s taxonomy.
3.4 Research Question 2: Can PBL Foster HOT and critical thinking?
To answer Research Question 2, the main findings on the effectiveness of PBL on creative thinking and HOT are summarized (see Table 2). First, the effects of PBL on critical thinking are described, followed by the four studies that examined HOT.

3.5 The effect of PBL on critical thinking
Based on the results of the analysis, the main findings of 27 studies investigating the effects of PBL on critical thinking were obtained. Five studies compared PBL with a control group and only reported posttest scores. The study under review used pre-experiment data from one PBL group (no control group). One study had another design (comparing groups; Pardamean, 2012). The most commonly reported control group was a control group (traditional learning). The study under review used pre-experiment data from one PBL group (no control group). One study had another design (comparing groups; Pardamean, 2012). The most commonly reported control group was a control group (traditional learning). The study under review used pre-experiment data from one PBL group (no control group). One study had another design (comparing groups; Pardamean, 2012).

Meta-analysis In the meta-analysis section of this review, we only included studies with pre-experiment and/or independent group designs. Independent group designs and preexperiment designs both provide insight into the question of whether PBL affects learners’ CT. The independent group design aims to test whether the learning method is more effective than “traditional” education, while the pre-experiment design examines differences before and after implementation. As mentioned, the impact is heterogeneous. Due to the limited number of studies, we could not investigate modifying factors that could statistically explain the heterogeneity. The variation in the magnitude of the effect may be partly due to variation in the way PBL is implemented, as this often differs across institutions even when the defense characteristics are met (Maudsley, 1999; Norman & Schmidt, 2000). However, to address this issue, we only included studies that met the definitional criteria of PBL and PjBL. We further excluded studies that contained additional activities (e.g. concept mapping) that could have influenced the results.

In addition, differences in the operationalization of CT may affect the results. To explore this, we calculated effect sizes separately for studies reporting outcomes on process and disposition. The analysis showed higher effect sizes for studies reporting critical thinking process outcomes (d=0.720, SE=0.14, 95% CI [0.46, 0.99]) than studies reporting critical thinking disposition effects (d=0.411, SE=0.13, 95% CI [0.16, 0.66]). However, these results should be interpreted with caution due to the limited number of studies and extreme effect sizes in the critical thinking process group.

Additional findings In the meta-analysis section of this review, we only included studies with a pre-experimental and/or independent group design. Two studies deviated from this design. Tiwari et al. (2006) not only compared the impact of PBL vs lecture-based learning but also included two follow-ups one and two years later. Results showed that the PBL group showed a significant increase in critical thinking scores, and the score increase remained positive at the first follow-up. However, the score gains became insignificant at the second follow-up (two years later). When we look at multiple subscale scores, the results show significant improvement in favor of PBL for truth-seeking, analytic, and CT confidence. The analytic score remained positive at the first follow-up, while the truth-seeking score remained significant at the first and second follow-up (two years later). The results of this study suggest that PBL can have long-term effects on critical thinking dispositions.

3.6 Effect of PBL on HOT
Only four of the twenty studies investigated the effect of PBL on HOT. Based on these four studies, it is not possible to calculate the effect size based on the data provided. Sugeng and Suryani (2020) compared a PBL group with a lecture-based group on HOT and low-level thinking. The PBL group scored significantly higher on HOT, while the lecture-based group scored higher on low-level thinking. Sasson et al. (2018) reported positive effects for a 2-year PBL program. HOT increased in the PBL group but not in the control group from Measurement 1 (beginning of grade 9) to Measurement 3 (end of grade 10).

4. Conclusion
This systematic review focuses on two questions: “How are HOT and critical thinking conceptualized in a learner-centered learning environment?” and “Can PBL foster HOT and critical thinking?” It has previously presented and discussed findings related to these questions. This section, therefore, offers a more global examination of the trends that emerged from the previous analysis of the
two forms of learner-centered learning approaches, PBL, and shares the remaining issues that should be explored in future research. However, it should be noted that several limitations emerged in this systematic review that affect the conclusions formulated. First, very limited attention has been given to HOT, with only four studies investigating it. There was also an uneven distribution in the studies, including on measures of thinking and learning environments. As mentioned, critical thinking measures used in these studies often consist of several subscales. For example, the WGCTA consists of subscales of inference, assumption recognition, deduction, interpretation, and argument evaluation. The presence of multiple indicators is a limitation as the impact on these sub-components can vary, making the interpretation of global effectiveness more difficult. However, in this study, it was not used as a search term in the literature search, and most studies did not report or define this subscale. Future research could use more detailed search terms, including subscales or process of thinking measures.

Although this study showed a positive effect of PBL on HOT and critical thinking, there are no in-depth details on what exactly caused the positive effect, as many relationships between PBL and HOT and critical thinking could not be identified. In addition, the relationship of HOT and critical thinking with performance was not investigated in this study. A more controlled experimental study could shed light on these issues and help overcome the design problems associated with effect studies. Finally, future research could link HOT critical thinking in a PBL environment with other learning processes, such as self-regulated learning (SRL) and self-directed learning (SDL). Components such as metacognition also play an important role in SRL and SDL processes. Future research can shed more light on the relationship between thinking and regulatory processes in the context of PBL.

5. Statement
In summary, this review research led to several conclusions regarding the conceptualization of HOT critical thinking in the PBL literature. First, HOT is not often studied in the PBL literature. Second, there is more focus on critical thinking skills (processes) than critical thinking dispositions in the research literature. Furthermore, the lack of clear conceptualization of HOT and critical thinking complicates measurement and findings. The lack of conceptual clarity leads to a limited number (and sometimes availability) of assessment instruments and tools and their questionable validity. The lack of conceptual clarity also extends to the PBL environment, where the essential components of PBL are not always articulated or discussed in studies that claim to apply this approach. Design issues in effects studies add to this complication. Furthermore, references to HOT or critical thinking skills.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers.

References


