Thabiea : Journal of Natural Science Teaching
Vol. 6(2), pp. 143-156, 2023
Available online at
http://journal.iainkudus.ac.id/index.php/Thabiea
p-issn: 2580-8474, e-issn: 2655-898X

Development instrument of interdisciplinary thinking skills in science lesson planning (IITSSL)

Santiani Santiani[1], Winarto Winarto*[2], Yatim Md Nadrah[3], Sri Jumini[4]
[1]Institut Agama Islam Negeri Palangkaraya, Jl. G. Obos, Menteng, Kec. Jekan Raya, Kota Palangka Raya,73112, Indonesia
[2]Universitas Negeri Yogyakarta, Jl. Colombo Yogyakarta No.1, Karang Malang, Caturtunggal, Kec. Depok, Daerah Istimewa Yogyakarta 5528, Indonesia
[3]Universitas Sains Islam Malaysia, Bandar Baru Nilai, Nilai, Negeri Sembilan, 71800, Malaysia
[4]Universitas Sains Al-Qur'an, Kemiri, Sukorejo, Kec. Mojotengah, Kabupaten Wonosobo, Jawa Tengah 56351, Indonesia
**Correspondence: winarto88@uny.ac.id*

Keywords:
Development; Instrument; Interdisciplina Skill; Thinking Skills; Science Lesson;

Abstract

The IITSSL is an interdisciplinary thinking assessment rubrics for science teacher candidates in a science lesson plan. The IITSSL rubric was developed and qualitative evaluated in three phases: rubric design, first and second pilot tests. The qualitative evaluation during three phases with novice-expert interviews was carried out in four science education classes at three different universities. The rubric development result of the IITSSL rubric with dimensions are objective, disciplinary grounding, integration, and critical awareness in a component science lesson plan are learning objectives, instructional activities, and assessments with valid categories of content validity and fair ICC (0.637) reliability. Evidence of novice and expert validity of the IITSSL rubric from the results of interviews with novices and experts, which in general are meaningful according to the values of validity and reliability. The IITSSL rubric with four dimensions and ten criteria as items is very simple, easy to use, and can in fact detect pupils' capacity for interdisciplinary thinking with accuracy and reliability. in science lesson plans in a range of course settings. The science teacher candidates appear to comprehend the importance of interdisciplinary thinking according to their experience of the science lesson plan coursework and the IITSSL rubric. The IITSSL rubric not only measure understanding of interdisciplinary thinking, but is likely to promote a more integrated method of knowledge how science teacher candidates solve real word problems

To cite this article:
Santiani, et al. (2023). Development instrument of interdisciplinary thinking skills in science lesson planning (IITSSL). *Thabiea : Journal of Natural Science Teaching*, 6(2), 143-156.

Introduction

Real life word problems are complex problems and must be solved from in several fields. Interdisciplinary thinking is a way to be able to solve problems from various perspectives (Zhang & Shen, 2015) (Suhodimtseva et al., 2020). Science can make life easier and can be a place for the formation of interdisciplinary thinking skills (Tripp & Shortlidge, 2019). Science learning is designed to be oriented towards interdisciplinary thinking skills (Cowden, 2016). Interdisciplinary learning must be complemented by an evaluation of interdisciplinary thinking skills in order to measure the achievement of the interdisciplinary learning carried out (Wang & Song, 2021).

Interdisciplinary thinking skills are high-level thinking skills with complex cognition (Spelt, 2017) and the ability to combine the expertise of two or more subjects (Mansilla & Duraisingh, 2007). Interdisciplinary thinking skills have dimensions such as disciplinary grounding, integration, critical awareness (Mansilla & Duraisingh, 2007), objectivity (Tripp & Shortlidge, 2020), humility, communication, and collaboration (Spelt, 2017) (Mansilla & Duraisingh, 2007). Performance assessment interdisciplinary thingking in science lesson planning (IDT-SLP) rubric for teacher candidates with dimensions of objectivity, disciplinary grounding, integration, and critical awareness.

The assessment can be used in interdisciplinary learning to shape and guide instruction (Versprille et al., 2017). Interdisciplinary science learning, research, and practice need reliable and valid instruments for assessment (Yang et al., 2018) (You et al., 2018). In higher education, there have been mostly analytic-descriptive rubrics for performance assessment (Brookhart, 2018). Performance assessment as an alternative assessment is part of high-quality curriculum materials and assessments (Hammond, 2012). Performance assessment with rubrics effective to assess the cycle of teaching (Brown, 2017), as a part of teaching approach (Koswara et al., 2021), and can to measure interdisciplinarily thingking in science (Tripp & Shortlidge, 2020).

Specific expectations for assignments in the form of scores as an assessment tool are rubrics (Stevens & Levi, 2005) that describe performances, complex student reasoning, or products (Arter, 2012). Rubrics are generally known by two parts: description and criteria (Stevens & Levi, 2005). Rubrics can be performance appraisals with detailed assessment descriptions and dimensions. Rubrics can be used in summative assessment, formative assessment, and instructional instruments (Jonsson & Svingby, 2007). The use of rubrics by students will lead them to instructional goals to achieve learning. Rubrics help students work on assignments according to assessment standards (Jonsson & Svingby, 2007). The IITSSLis a rubric for assessing the interdisciplinary thinking of prospective science teacher candidates in science learning plans.

The science lesson plan is the main support for the success of science class management in improving the quality of learning and assessment (Iqbal et al., 2021). There are still many problems faced by teachers in preparing lesson plans. The teacher still experienced difficulties in developing authentic mathematics assessments (Pardimin, 2018). The student teacher candidates must enhance their PCK dimension in the lesson plan for teaching science (Maryati et al., 2019). Instruments that are valid, reliable, and suitable for assessing lesson plans must be developed. The IITSSL rubric is used to assess the interdisciplinary thinking skills of science teacher candidates in the science lesson plans, begins by outlining the learning goals, then constructing instructional activities, and creating evaluations (John, 2015) as a key to lesson planning (Chizhik & Chizhik, 2018). The IITSSL rubric is designed by integrating the dimensions of interdisciplinary thinking into the lesson plan component.

The role of novice-expert studies is defining the educational learning outcome and the learning goal (Schunn & Patchan, 2009). Expert-novice studies examine the natural disparities between people performing at relatively levels of a specific domain (Kamarudheen, 2015). The IITSSL rubric uses novice-expert studies, with novices being science teacher candidates and experts being lecturers or instructors. The instrument's observed scores and the pupils' responses are inextricably linked (Wren & Barbera, 2013). This procedure can help instrument designers make sure that students understand the criteria as intended and also give them a clearer view of the thought processes involved in creating a student's response. The technical term for this is beginner response process validity (AERA, 2014). The expert response process is a method for gathering more information on whether the scoring scale is appropriate from subject-matter experts (AERA, 2014).

The experts have verified the content validity of a number of rubrics (Jonsson & Svingby, 2007). The five primary types of construct validity evidence are test content, response process, internal structure, relationship with other variables, and consequence of use (AERA, 2014). Numerous techniques, including factor analysis and expert reviews, were used to examine construct validity; some research additionally addressed consequential evidence for validity using questionnaires given to students or teachers (Brookhart, 2018). This study uses content validity by expert jugment and construct validity by novice-expert response process.

Consistency among raters is crucial for determining the caliber of student performance. The Cohen's kappa, intra-class correlation coefficient (ICC), and percentage of rater agreement are often used metrics to assess the reliability of a rubric (Brookhart & Chen, 2015). The ICC is a widely used repeatability metric for continuous data sets that includes a variety of versions that measures the degree of correlation and agreement between measurements (Perinetti, 2018).

The ability to obtain evidence of validity is necessary for the assessment and evaluation tools (Paul et al., 2019). The IITSSL rubric used evidence validity with novice and *expert* response processes. The novice response processes help instrument developer ensure that students are interpreting the indicator correctly or according to how they respond. experts on whether the grading system is appropriate and the criteria within the constructs are able to interpret the results as intended as part of the expert response process.

Preparing students to produce quality interdisciplinary work needs assessment (Mansilla & Duraisingh, 2007). Performance assessment with rubrics interdisciplinary thingking for students teacher candidates very limited. Only two research articles were found that validated the interdisciplinary thinking performance assessment with rubrics for assessing students' essays or writing science tasks (Mansilla & Duraisingh, 2007 ; Tripp & Shortlidge, 2020). The development of an interdisciplinary thinking performance assessment instrument with rubrics for student teacher candidates preparing a science lesson plan has never existed. Interdisciplinary thinking skills for student teacher candidates must be developed to solve life problems and develop environmentally-oriented science interdisciplinary learning. Performance assessment using valid and reliable interdisciplinary thinking rubrics must be developed. What evidence from novice-expert response (qualitatif evaluation) support the constructs a quality of the IITSSL rubric ?

Method

The research question was answered through the development and two pilot tests of the IITSSL rubric and semi-structured student and faculty interviews (Figure 1). The IITSSL rubric was developed in three phases: rubric design, first rubric pilot, and second rubric pilot (Tripp

& Shortlidge, 2020). The four processes of developing rubrics for the IITSSL rubric dimensions are reflecting, listing, grouping and labeling, and application (Stevens & Levi, 2005).

Rubric Development : Phase 1

The reflecting stage of rubric design determines the purpose of compiling the rubric and taking dimensions based on the literature (Stevens & Levi, 2005). We would like to use the term "interdisciplinary thinking," not "interdisciplinary understanding" (Tripp & Shortlidge, 2020). The IITSSL rubric has four dimensions: objective, disciplinary grounding, integration (Tripp & Shortlidge, 2020), and critical awareness (Mansilla et al., 2009). The lesson plan components are learning objectives, instructional activities, and assessments (John, 2015; Chizhik & Chizhik, 2018) as part to be graded with the IITSSL rubric.

The rubric design at the listing stage focuses on the specifics of the task and writing of learning objectives (LOs) for the assignment (Tripp & Shortlidge, 2020 ; Stevens & Levi, 2005). The LO of the IITSSL rubric is the preparation of an interdisciplinary science lesson plan with a determination stage objective (LO1), disciplinary grounding (LO2), integration (LO3), and critical awareness (LO4). In stage 3, we grouped the results of our reflections and listed them, grouping and labeling similar objectives to become the rubric dimensions and identifying criteria that would define each construct (Tripp & Shortlidge, 2020 ; Stevens & Levi, 2005). LO1 has three criteria: purpose, approach, and credibility (Tripp & Shortlidge, 2020) for the science lesson plan component. LO2 has three criteria: disciplines, disciplinary reasoning, and method and tool (Tripp & Shortlidge, 2020) for the instructional activities of the science lesson plan component. LO3 has two criteria: leveraging disciplines and collaboration (Tripp & Shortlidge, 2020) for the instructional activities of the science lesson plan component. LO4 (Mansilla & Duraisingh, 2007) with two criteria, societal impact and limitations (Tripp & Shortlidge, 2020), for the assessment of the science lesson plan component The last stage, compiling a grid format, included using the constructs and relevant criteria. LO1 with three items, LO2 with three items, LO3 with four items, and LO4 with two items.

The Rubric's Level or Scale

A scoring rubric only includes one set of standards to gauge how pupils have responded to the rubric's dimensions (Tripp & Shortlidge, 2020 ; Stevens & Levi, 2005). The IITSSL rubric uses four criteria: naive (1), novice (2), intermediate (3), and master (4) (Tripp & Shortlidge, 2020 ; Mansilla et al., 2009). Then, for each construct domain of the lesson plan that they have designed, we established criteria that would characterize science teacher candidates' ability to think interdisciplinary. Though not intended to be construed as more or less significant, certain conceptions contained more criteria than others. Using the rubric, lecturer would average the results of each construct's criterion to provide a single score for each construct.

Science Lesson Plan Assigment

We gathered samples of student teacher candidates' science lesson plans that allowed them to demonstrate interdisciplinary thinking in order to test the IITSSL rubric. In our earlier work, we created course-specific lesson plan assignments that required students to combine their knowledge from science and peatland conservation fields to solve a complex real-world problem. Here, we applied the same science lesson plan structure to construct fresh, pertinent prompts that we worked on with each of the study's course instructors. The students were provided the "student version" of the IITSSLproject requirements rubric to use as a self-assessment the scales used to grade novice (mastery, intermediate, novice, and naive). It was similar to the complete "practitioner version" for instructors or experts.

First Pilot Test of Rubric : Phase 2

Recruitment and Data Collection

The preliminary iteration of the IITSSL rubric and and a related task for a science education course (A1) at Palangka Raya State Islamic Institute Religion (IAIN Palangka Raya) for three months (August–October) in the odd semester of 2022 (Table 1). In this course, students are provided with five meetings to develop lesson plans with interdisciplinary learning of science and peatland conservation with researcher 1 as instructor. We were able to use the replies to the lesson plan assignment for this study since students in this course had given their approval for any course material to be used for research.

Table 1. Sample sizes of science lesson plan and interviews collected in three universities over the course of one semester (D, disciplinary; STC, Science Teacher Candidates ; SLP, Science Lesson Plan ; II, Instructor Interviews)

University	Course Department : Format	SLP (n)	STC (n)	II (n)
A1 : Teaching practice [a]	Physic Education : D	10	5	-
A2 : Physic Lesson Plan[a]	Physic Education : D	12	5	1
B : Science learning development for elementary schools [b]	Elementary Education : D	27	5	1
C : Microteaching [c]	Physic Education : D	10	4	1
Total		49	19	3

[*a]A university, [b] B university, [c] C University

As part of the science lesson plan assignment, the instructor distributed the student version of the IITSSL rubric. At the end of course after five meetings, were given approximately seven days to complete the individual science lesson plan assignment with interdisciplinary science and peatland conservation to science teacher candidates. When the lesson plan complete, we conducted an self-assessment process using the student version of the IITSSL rubric and semi-structure interview with form of interview. The instructor also conducts assessments using the IITSSL rubric.

Interviews with Science Teacher Candidates

We performed think-aloud interviews with novice (science teacher candidates) that were semi-structured to better understand how they were understanding about the construct in the IITSSL rubric in order to show evidence the validity of novice response process. Three researchers developed interview questions and iterated on them, to know how the categories and criteria were phrased had an impact on how students responded. The science teacher candidates understood the rubric as we had meant. How they felt about the assignment and rubric, as well as any additional ways they might have understood interdisciplinary science besides those covered by the rubric. In order to find evidence that students understood the requirements in the IITSSL rubric as we intended, we deductively studied interview transcripts for the IITSSL rubric (Tripp & Shortlidge, 2020). We did not interview with the course instructor because she is the author of this paper.

Content validity and evidence novice validity of The IDT-SLP

The IITSSL rubric and items were reviewed by five experts in science education at Palangka Raya University. The experts are practitioners and researchers of environmental-based science education, especially peatland conservation in Palangka Raya. Reviewed five experts led the validation process and calculated content validity by Aiken´s value (Aiken, 1985; Indriastuti et al., 2020; Luque et al., 2018).

The score with the IITSSL rubric were graded by researchers as instructur in first pilot test. We then collaborated to determine how to accurately reflect the goals of each criterion and provide clarity to sections that may have been unclear or potentially deceptive to students. Based on student interviews and answers to scientific lesson plans from the first pilot test, the rubric was modified in some areas for the second test.

Second Pilot Testing of Rubric : Phase 3

Recruitment and information gathering

We piloted the IITSSL rubric and science lesson plan in phase 3 assignments will be three more courses at different universities, testing the rubric's suitability for different demographics. One course from the same university (courses A2), and two courses from two separate universities (courses B and C). Course A2 is a science lesson planning course at the same university as A1. Course B is a science learning development for elementary schools at University B, which is different from University A. Course C is a microteaching course at a different university from A and B (Table 1). A schematic outline of development and qualitatif evaluation of the rubric is visualized in Figure 1.

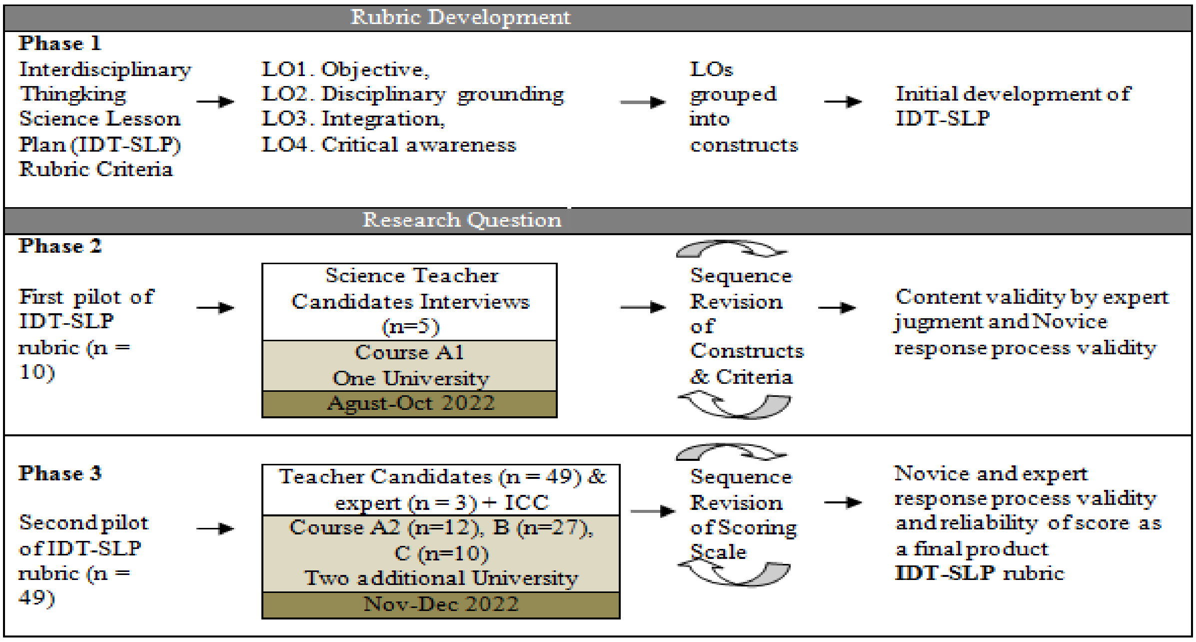

Figure 1. Phases 1, 2, and 3 of the IITSSL rubric's development and qualitatif evaluation (LO, learning objective; ICC, interclass correlation coefficient)

Courses A, B, and C are not interdisciplinary courses, so before carrying out the assignment and assessment, lectures were carried out five meetings with interdisciplinary learning to shape students' abilities to prepare science lesson plans. The data collection process in phase 3 is the same as in phase 2. After five meetings, students develop a science lesson plan with interdisciplinary science and peatland conservation. The science lesson plans that have been prepared are self-assessed by the students using the IITSSL rubric and reported to the instructor. Students then fill out the interview form as a response to the IITSSL rubric that has been used.

Phase 3 is carried out at the same time in October-December, the odd semester of 2022, with interdisciplinary science and peatland learning with instructors who are not researchers. The instructors in the second pilot test are lectureas expert in science education based on peatlands environmental. We were able to evaluate the IITSSL rubric's functionality irrespective of the population, student major, or course format by applying the IITSSL rubric criteria to a variety of interdisciplinary thinking formats from various institution kinds. The student interviews were conducted in the same semi-structured, think-aloud manner as in phase 2, and the same interview questions were used.

Faculty interviews were conducted through semi-structured interviews with the instructors (n = 3) after the classes had ended. We were able to acquire information and feedback on the rubric and assignment, and this allowed us to demonstrate the validity of the expert response process. A semi-structured interview sheet was utilized to get the teachers' permission to use their interview responses after we contacted them and requested their participation. Three researchers developed and iterated on interview questions to learn more about the functionality and the validity of each concept and criterion of the rubric's instructor version.

Interclass Correlation Coefficient (ICC) Reliability

One week before the interviews, the researcher contacted each instructor and asked them to grade the scientific lesson plans from their respective classes using the IITSSL rubric. We used ICC values to examine the reliability of two raters (instructors) who evaluated a continuous data set with the IITSSL rubric. According to the ICC, the following thresholds were used: under 0.50, bad; 0.50 until 0.75, fair; 0.75 until 0.90, good; and above 0.90, excellent (Perinetti, 2018). These reliability analyses were carried out using SPSS 18.

We used a ANOVA with SPSS 18 to find variations in the mean student essay scores between courses A2–C (Schunn & Patchan, 2009) in order to evaluate the usefulness of the updated rubric across various populations from phase 3. As we repeatedly refined the rubric to be generally relevant to any discipline and actual situations challenge, we hypothesized that there wouldn't no variation in a comparison of courses' overall essay scores. Due to the difficulty in finding specialists in significant numbers, many statistical tests' minimum N assumptions are not met, necessitating the usage of low-N versions of those tests (Kamarudheen, 2015).

Results and Discussion

Dimensions of IITSSL rubric

The final rubric uses four constructs and related criteria to evaluate how well science teacher candidates apply interdisciplinary thinking to develop a science lesson plan. The dimensions are objectivity, disciplinary grounding, integration, and critical awareness. This was determined by reviewing interview transcripts from science teacher candidates (n = 5) in phase 2. The rubric's four constructs from phase two were carried over to phase three. We have designated these constructs as "categories" to help instructors who aren't familiar with the term "construct" comprehend this specific rubric dimension.

Evidence of Validity from Novice-Expert Response Process

For each criterion in the IITSSL rubric, we have supplied teacher candidates' science lesson plan and interview replies, as well as faculty interview responses, in Table 2. Below, we outline more justifications for including or excluding each criterion from our instrument's piloting.

Table 2. Examples of student essays, faculty interview responses, and interview responses from students that support the IITSSLcriterion (STC, Science Teacher Candidates; SLP, Science Lesson Plan)

No	Construct	Criteria	Example of STC of SLP from course B	STC interviews from all courses	Instructor Interview
1	Objective	Purpose	Causes, impacts, and prevention of global warming with peatland conservation	“Helps and makes it easier to choose problems in SLP”-course A1	“STC determine the problem at this stage of the goal”
		Approach	The interaction of living things in the peatland environment	“Help in setting learning objective”-course A2	” This rubric helps STC in choosing learning outcomes
		Credibility	Peat water fluid dynamics	“Be a guide in writing complete learning objectives”- course C	“STC can complete the learning objective”
2	Disciplinary Grounding	Disciplines	Peatland environmental pollution	“Determine the material science and conservation of peatlands”-course B	“Make it easier for students to choose teaching materials”
		Disciplinary reasoning	Application of electromagnetic waves to measure the depth of peatlands	“Using peatland conservation technology in learning”-course C	“STC choose peatland conservation technology with a suitable pedagogical approach”
		Methods and tools	Peatland Thermodynamic	“Using scientific experimental methods on peatlands”-course A2	“take the peatland scientific or philosophical experimental method”
3	Integration	Leveranging disciplines	Classification of matter and its changes	“Using problem based learning model”-course A1	“Students can choose the appropriate learning model and strategy”
		Collaboration	The nature of the continuity of peat water and ordinary water	“Using learning media by integrating science and peatlands”-course A2	“Students can combine science and peatland material in compiling learning media”

4	Critical Awareness	Limitations	Peatland plant tissue system	“Cognitive assessment using a combination of science and peatlands with not standarized”-course C	"STC composes cognitive assessments using a combination of science and peatlands, but it is not a valid instrument."
		Societal impact	Technology based on plant tissue systems (physical peat water filtration)	"Assess the attitude of caring for students' peatlands; it is not valid." -- Course A2.	"Choose an attitude assessment that fits the learning objectives but is not standard."

Objective Construct

According to the opinion of the five experts who assessed the content validity of the IITSSL rubric, the objective dimension with the criteria of purpose, approach, and credibility has an Aiken value in the valid category. The objective dimension in the initial development consisted of nine items. In the content validation stage by experts, there are 6 items with a valid Aikens value of 0.87. In the second phase, we only used three items with consideration of simplifying the rubric with input from the experts.

The objective dimension is generally considered by students and teacher candidates to be the easiest dimension to construct. This dimension and its criteria are very helpful in preparing an interdisciplinary science lesson plan for science and peatlands (Tripp & Shortlidge, 2020). The objective aspect examines the extent to which students articulate their interdisciplinary writing goals and audience (Mansilla et al., 2009).

Three criteria on the objective dimension can be very well determined by students. Interviews with instructors who have used the interdisciplinary thinking performance assessment rubric also state that the objective dimension is the easiest to determine and helps students develop interdisciplinary learning objectives. The learning objectives in the science lesson plans are described with the aim of incorporating the dimensions of interdisciplinary thinking (Table 2), so that the prepared lesson plans are interdisciplinary lesson plans. Conformity of the opinions of students and instructors can be seen in Table 2.

Disciplinary Grounding

The instructional activity component of the science lesson plan contains a disciplinary dimension grounded in three criteria: discipline, discipline reasoning, and methods and tools. According to the five experts, this dimension has the Aikens value of 1 with a valid category for three items. This validity is supported by evidence from student and instructor responses obtained through interviews, which state that the criteria of discipline help in selecting disciplines related to the problems that have been determined.

The responses of students and lecturers related to this rubric focus on the criteria of disciplinary reasoning, namely helping students choose technology that will be associated with learning materials, namely science technology and peatland conservation. This result is in line with the purpose of this dimension, which is to examine students' comprehension, choice, and application of the bodies of knowledge that influence their job (Mansilla et al., 2009).

However, new students are able to choose technologies, methods, and tools related to scientific disciplines, not peatland conservation technology. The instructor's response stated that in general, students did not use integrated technologies, methods, and tools from peatland conservation and science. This happens due to a lack of student knowledge about other disciplines (Tripp & Shortlidge, 2020). That enabled them to articulate two or more disciplinary understandings (Mansilla & Duraisingh, 2007) through learning interdisciplinary design (Nae, 2017; Suhodimtseva et al., 2020), interdisciplinary problem solving (Zhang & Shen, 2015), or an interdisciplinary modul (Cowden, 2016).

Integration Constructs

Integration dimension with two criteria are leveraging discipline and collaboration with content validity according to the Aiken score of five raters categorized as valid (0.9). Evidence of the validity of this content can be seen in the results of interviews with students and instructors who have the same meaning, namely, that the rubric is very helpful in integrating between disciplines.

In the second phase of class A1, students only register models, strategies, and learning media from different disciplines and have not integrated them. To improve students' integration abilities before the meeting in the third phase, the researcher presents examples of integration that has been carried out with the key words of integration. In the third phase, in classes A2, B, and C, students begin to choose a learning model that is appropriate to the disciplines of science and peatland conservation. The media used in the science lesson plan design is an integration of the two disciplines.

According to students, the collaboration criteria were the most difficult to meet. Interdisciplinarity between the natural sciences and social sciences is quite difficult due to epistemological differences or philosophical problems between different disciplines (Graff, 2016; Tripp & Shortlidge, 2019). Collaboration is the core criterion of interdisciplinary thinking. The ability of students to combine viewpoints is the definition of interdisciplinary understandings (Mansilla & Duraisingh, 2007). As an initial stage, forming interdisciplinary thinking skills is an integration dimension that involves creating student learning teams from various disciplines through cooperative learning (Harvie, 2012). Incorporating local community knowledge into learning can be a pragmatic way of training students' integration skills.

Critical Awareness Construct

Limitation and societal impact are the two requirements for the critical awareness dimension, which is an openly stated meta-disciplinary stance on their interdisciplinary activity (Mansilla et al., 2009). Assessment in the science lesson plan with the criteria, limitations, and societal impact Content validity as determined by the five raters or experts on this dimension is categorized as valid based on the Aiken value (0.9). Evidence for the content validity of this dimension, as shown in Table 2, shows similar positive responses from students and instructors. The assessment is based on interdisciplinary knowledge of science and peatland conservation as a criterion limitation. The attitude that will be formed is related to environmental literacy as a social impact criterion. The final response of students and instructors after using the IITSSL rubric is a meta-analysis of the science lesson plan that has been prepared. After going through this stage, the two parties agreed that there were many limitations to the lesson plan that they had compiled and that further development research should be carried out.

The science teacher candidates appear to comprehend the importance of interdisciplinary thinking according to their experience of the science lesson plan coursework and the IITSSL rubric. The IITSSL rubric not only measure understanding of interdisciplinary thinking, but is likely to promote a more integrated method of knowledge how science teacher candidates solve real word problems.

Interclass Correlation Coefficient (ICC) Reliability, and Statistical Findings

The score interpretations of two instructor from lecturer (never used our instrument before and lacked any training) and researcher have ICC (0.637) as fair reliability. This demonstrates that users of the IITSSL rubric can apply it without having received any formal training in how to evaluate or interpret student work, and it also demonstrates the validity of the information gathered using the IITSSL rubric (Tripp & Shortlidge, 2020).

In examining differences in students' interdisciplinary thinking across populations, we found differences in the mean score of the science lesson plan in courses A2, B, and C ($F = 3.724$, $p = 0.03$). ANOVA test followed by the Pos Hoc test to see the F value of each class ($F_{A2} = 1.353$, $F_B = 2161.814$, and $F_C = 1.353$). Pos Hoc test results show that A2 and C courses have mean values that are not varied or relatively the same, while class B average lesson plan scores are significantly different ($F_B = 2161.814$). In general, the score of the science lesson plan with the IITSSL rubric in course B was not too much different, but the ANOVA analysis was influenced by the number of students in course B ($n = 27$), which was far more than the other courses (A2 = 12 and C = 10).

We used SPSS 18 to conduct a post hoc power analysis based on the small sample sizes at the course level (Table 1), which showed the necessity for bigger sample sizes to accurately assess variations in students' ID thinking across populations and institutions (Tripp & Shortlidge, 2020). The rubric developed is an alternative assessment that is qualitative in nature, so it is quite difficult to manage large amounts of data. The small number of experts also makes it difficult to increase the sample size (Schunn & Patchan, 2009; Kamarudheen. K, 2015). But as The IITSSL rubric can definitely identify students' capacity to think interdisciplinaryly in scientific lesson plans in a range of course settings, as demonstrated by the validity and reliability tests from novice-expert interviews (Tripp & Shortlidge, 2020). Course A1 was excluded from our statistical analysis since it was given the original IITSSL rubric (in phase 2) and is therefore statistically incompatible with the other courses. The validity and reliability tests provide evidence that the IITSSL rubric may accurately and consistently identify students' capacity for interdisciplinarity in scientific lesson plans across a range of course settings.

Conclusion

The IITSSL rubric is a student interdisciplinary thinking assessment rubric for science teacher candidates in a science lesson plan. The IITSSL rubric was developed and qualitatively evaluated in three phases: rubric design, the first pilot test, and the second pilot test. The qualitatif evaluation during three phase with novice-expert interviews. Four processes of developing rubrics design are reflecting, listing, grouping and labeling, and application. Rubric development results in the IITSSL rubric having four dimensions: objectives, disciplinary grounding, integration, and critical awareness. In a component science lesson plan, these are learning objectives, instructional activities, and assessments. The implications of this research are that interdisciplinary thinking skills can be measured using the lesson plan that was designed. Therefore, developing interdisciplinary thinking skills can be evaluated from the lesson plan that is prepared. The advantage of this research is that the non-test instrument helps lecturers identify the level of interdisciplinary thinking skills of prospective teachers. The weakness is that the test subjects in developing this instrument were less than one hundred subjects

Acknowledgment
Thanks to Prof. Agus Haryono, Palangkaraya State University, who has been involved and contributed to this research

Credit Authorship Contribution Statement

Santiani constructing and reviewing the literature. **Nadrah** and **Sri Jumini** reviewed the literature and edited the manuscript by **Winarto**. All authors read and approve the final manuscript.

References

Aiken, L. R. (1985). Three Coefficients for Analyzing the Reliability and Validity of Ratings. *Educational and Psychological Measurement*, 45(1), 131–142. https://doi.org/10.1177/0013164485451012

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (Eds.). (2014). *STANDARDS f o r E d u c a t i o n a l a n d Ps y c h o l o g i c a l Te s t i n g*. American Educational Research Association.

Arter, J. (2012). *Creating & Recognizing Quality Rubrics a Study Guide from Pearson Assessment Training Institute*. Pearson Assessment Training Institute, Portland, Oregon, www.ati.pearson.com

Mansilla, Veronica., & Duraisingh, E. Dawes. (2007). Targeted Assessment of Students' Interdisciplinary Work: An Empirically Grounded Framework Proposed. *The Journal of Higher Education*, *78*(2), 215–237. https://doi.org/10.1353/jhe.2007.0008

Brookhart, S., & Chen, F. (2015). The quality and effectiveness of descriptive rubrics. *Educational Review*, *Query date: 2022-10-27 16:19:48*. https://doi.org/10.1080/00131911.2014.929565

Brookhart, S. M. (2018). Appropriate Criteria: Key to Effective Rubrics. *Frontiers in Education*, *3*, 22. https://doi.org/10.3389/feduc.2018.00022

Carmen Sherry Brown. (2017). *Iligning a Performance-based Observation Rubric to Support a Teacher Performance Assessment*. *5*(2), 15

Chizhik, E. W., & Chizhik, A. W. (2018). Using Activity Theory to Examine How Teachers' Lesson Plans Meet Students' Learning Needs. *The Teacher Educator*, *53*(1), 67–85. https://doi.org/10.1080/08878730.2017.1296913

Cowden, C. (2016). Interdisciplinary Explorations: Promoting Critical Thinking via Problem-Based Learning in an Advanced Biochemistry Class. *Journal of Chemical Education*, *93*(3), 464–469. https://doi.org/10.1021/acs.jchemed.5b00378

Darling-Hammond, L. (2012). *Creating a Comprehensive System for Evaluating and Supporting Effective Teaching*. Stanford Center for Opportunity Policy in Education

Graff, H. J. (2016). The "Problem" of Interdisciplinarity in Theory, Practice, and History. *Social Science History*, *40*(4), 775–803. https://doi.org/10.1017/ssh.2016.31

Harvie, J. (2012). *Interdisciplinary Education and Co-operative Learning: Perfect Shipmates to Sail against the Rising Tide of 'Learnification'?*

Indriastuti, N., Sugini, & Anwar, M. (2020). Visually Impaireds Critical Thinking Skills (A Comparative Study between Inclusive School and Special School). *Proceedings of the 4th International Conference on Learning Innovation and Quality Education*, 1–5. https://doi.org/10.1145/3452144.3453764

Iqbal, Md. H., Siddiqie, S. A., & Mazid, Md. A. (2021). Rethinking theories of lesson plan for effective teaching and learning. *Social Sciences & Humanities Open*, *4*(1), 100172. https://doi.org/10.1016/j.ssaho.2021.100172

John, Y. (2015). A" New" Thematic, Integrated Curriculum for Primary Schools of Trinidad and Tobago: A Paradigm Shift. *International Journal of Higher Education*, *Query date: 2022-10-27 16:19:48*. https://eric.ed.gov/?id=EJ1088730

Jonsson, A., & Svingby, G. (2007). The use of scoring rubrics: Reliability, validity and educational consequences. *Educational Research Review*, *2*(2), 130–144. https://doi.org/10.1016/j.edurev.2007.05.002

Kamarudheen. K, B. (2015). Role of Expert Novice Studies in Education Research. *SRJIS*, *VOL-3*(21)

Koswara, D., Dallyono, R., Suherman, A., & Hyangsewu, P. (2021). The analytical scoring assessment usage to examine Sundanese students' performance in writing descriptive texts. *Jurnal Cakrawala Pendidikan*, *40*(3), 573–583. https://doi.org/10.21831/cp.v40i3.40948

Mansilla, V. B., Duraisingh, E. D., Haynes, C., & Wolfe, C. R. (2009). Targeted Assessment Rubric: An Empirically Grounded Rubric for Interdisciplinary Writing. *The Journal of Higher Education*, *80*(3), 21

Maryati, M., Prasetyo, Z. K., Wilujeng, I., & Sumintono, B. (2019). Measuring Teachers' Pedagogical Content Knowledge Using Many-Facet Rasch Model. *Jurnal Cakrawala Pendidikan*, *38*(3), 452–464. https://doi.org/10.21831/cp.v38i3.26598

Nae, H.-J. (2017). An Interdisciplinary Design Education Framework. *The Design Journal*, *20*(sup1), S835–S847. https://doi.org/10.1080/14606925.2017.1353030

Pardimin, P. (2018). Analysis of the Indonesia Mathematics Teachers' Ability in Applying Authentic Assessment. *Cakrawala Pendidikan*, *37*(2), 170–181

Paul, C. R., Ryan, M. S., Dallaghan, G. L. B., Jirasevijinda, T., Quigley, P. D., Hanson, J. L., Khidir, A. M., Petershack, J., Jackson, J., Tewksbury, L., & Rocha, M. E. M. (2019). Collecting Validity Evidence: A Hands-on Workshop for Medical Education Assessment Instruments. *MedEdPORTAL*, 10817. https://doi.org/10.15766/mep_2374-8265.10817

Perinetti, G. (2018). StaTips Part IV: Selection, interpretation and reporting of the intraclass correlation coefficient. *South European Journal of Orthodontics and Dentofacial Research*, *5*(1). https://doi.org/10.5937/sejodr5-17434

Schunn, C., & Patchan, M. (2009). *Expert-novice Studies: An Educational Perspective.*

Spelt, E. J. H. (2017). A multidimensional approach to examine student interdisciplinary learning in science and engineering in higher education. *European Journal of Engineering Education*, *42*(6), 761–774. https://doi.org/10.1080/03043797.2016.1224228

Stevens, D. D., & Levi, A. (2005). *Introduction to rubrics: An assessment tool to save grading time, convey effective feedback, and promote student learning / Dannelle D. Stevens, Antonia Levi* (1st ed). Stylus Pub

Suhodimtseva, A. P., Vorozheikina, N. I., & Eremina, J. B. (2020). Integration Approach to Solving Problems of Interdisciplinary Nature in the Conditions of Post-industrial

Education. In D. B. Solovev (Ed.), *Smart Technologies and Innovations in Design for Control of Technological Processes and Objects: Economy and Production* (Vol. 138, pp. 501–510). Springer International Publishing. https://doi.org/10.1007/978-3-030-15577-3_48

Torres-Luque, G., Fernández-García, Á. I., Cabello-Manrique, D., Giménez-Egido, J. M., & Ortega-Toro, E. (2018). Design and Validation of an Observational Instrument for the Technical-Tactical Actions in Singles Tennis. *Frontiers in Psychology*, *9*, 2418. https://doi.org/10.3389/fpsyg.2018.02418

Tripp, B., & Shortlidge, E. (2020). From theory to practice: Gathering evidence for the validity of data collected with the Interdisciplinary Science Rubric (IDSR). *CBE—Life Sciences Education*, *Query date: 2022-10-27 16:19:48*. https://doi.org/10.1187/cbe.20-02-0035

Tripp, B., & Shortlidge, E. E. (2019). A Framework to Guide Undergraduate Education in Interdisciplinary Science. *CBE—Life Sciences Education*, *18*(2), es3. https://doi.org/10.1187/cbe.18-11-0226

Versprille, A., Zabih, A., Holme, T. A., McKenzie, L., Mahaffy, P., Martin, B., & Towns, M. (2017). Assessing Student Knowledge of Chemistry and Climate Science Concepts Associated with Climate Change: Resources To Inform Teaching and Learning. *Journal of Chemical Education*, *94*(4), 407–417. https://doi.org/10.1021/acs.jchemed.6b00759

Wang, Z., & Song, G. (2021). Towards an assessment of students' interdisciplinary competence in middle school science. *International Journal of Science Education*, *43*(5), 693–716. https://doi.org/10.1080/09500693.2021.1877849

Wren, D., & Barbera, J. (2013). Gathering Evidence for Validity during the Design, Development, and Qualitative Evaluation of Thermochemistry Concept Inventory Items. *Journal of Chemical Education*, *90*(12), 1590–1601. https://doi.org/10.1021/ed400384g

Yang, Y., Peng He, & Xiufeng Liu. (2018). Validation of an Instrument for Measuring Students' Understanding of Interdisciplinary Science in Grades 4-8 over Multiple Semesters: A Rasch Measurement Study. *International Journal of Science and Mathematics Education*, *16*(4), 639–654. https://doi.org/10.1007/s10763-017-9805-7

You, H., Marshall, J., & Delgado, C. (2018). Assessing students' disciplinary and interdisciplinary understanding of global carbon cycling. *Journal of Research in …*, *Query date: 2022-10-27 16:19:48*. https://doi.org/10.1002/tea.21423

Zhang, D., & Shen, J. (2015). Disciplinary Foundations for Solving Interdisciplinary Scientific Problems. *International Journal of Science Education*, *37*(15), 2555–2576. https://doi.org/10.1080/09500693.2015.1085658