Prisma Sains: Jurnal Pengkajian Ilmu dan Pembelajaran Matematika dan IPA IKIP Mataram
https://e-journal.undikma.ac.id/index.php/prismasains/index
e-mail: prismasains.pkpsm@gmail.com

October 2026. Vol. 14, No. 4
p-ISSN: 2338-4530
e-ISSN: 2540-7899
pp. 2592-2606

Deep Learning with Dayak Onion–Ginger Ethnochemistry: Changes in Environmental Awareness and Green Chemistry Understanding in Two Intact Classes

[1]*Anna Farida, [2]Dyah Worowirastri Ekowati, [3]Agus Tinus
Master of Pedagogy Program, Postgraduate Directorate, Universitas Muhammadiyah Malang, Malang, Indonesia

*Corresponding author e-mail: annafarida44@guru.sma.belajar.id

Received: September 2026; Revised: October 2026; Published: October 2026

Abstract

Green chemistry is often taught in abstract terms, and locally relevant contexts may help students connect its principles to everyday practice. This preliminary study describes learning outcomes after Deep Learning lessons built around a Dayak onion (Eleutherine palmifolia) and ginger (Zingiber officinale) infusion. Two intact tenth-grade classes at SMAN 1 Sampit, Central Kalimantan, took part (36 students each, N = 72) in a pretest–posttest design with non-equivalent groups. One class received the Deep Learning lessons and the other received conventional instruction. Environmental awareness was operationalised as ecological worldview and measured with a locally adapted New Ecological Paradigm (NEP) questionnaire (maximum 75). Understanding of green chemistry was measured with a 20-item two-tier test (maximum 20). Mean awareness rose from 48.83 to 64.56 in the experimental class and from 49.28 to 54.00 in the control class; mean concept scores rose from 5.75 to 14.33 and from 5.67 to 9.17. Averaged over students, N-Gain was 0.59 versus 0.19 for awareness and 0.62 versus 0.27 for understanding. The findings are descriptive and exploratory, and no inferential test was applied.

Keywords: Deep learning; Ethnochemistry; Green chemistry; Environmental awareness; Dayak onion; Ginger.

How to Cite: Farida, A., Ekowati, D. W., & Tinus, A. (2026). Deep Learning with Dayak Onion–Ginger Ethnochemistry: Changes in Environmental Awareness and Green Chemistry Understanding in Two Intact Classes. *Prisma Sains: Jurnal Pengkajian Ilmu Dan Pembelajaran Matematika Dan IPA IKIP Mataram*, *14*(4), 2592–2606. https://doi.org/10.33394/j-ps.v14i4.22776

https://doi.org/10.33394/j-ps.v14i4.22776

Copyright© 2026, Farida et al.
This is an open-access article under the CC-BY License

INTRODUCTION

Twenty-first-century schooling is expected to do more than pass on content. It should help students reason critically, act on what they know, and stay alert to problems in their own surroundings Indonesia's Ministry of Primary and Secondary Education has responded by promoting Deep Learning, an approach built around mindful, meaningful, and joyful learning experiences (Kementerian Pendidikan Dasar dan Menengah, 2025). Deep learning is commonly distinguished from surface learning, which aims at success in tests rather than at understanding that can be applied in new situations (Fullan et al., 2018). In chemistry, the approach is presented as a way to address low scientific literacy and the tendency of students to hold concepts in isolation from one another (Fatihah & Wiji, 2026). The need for such approaches is documented in national and international assessment. In PISA 2022, Indonesian students scored 383 points in science against an OECD average of 485 (OECD, 2023). A systematic review of Indonesian studies grouped the difficulties of learning chemistry into internal factors, such as low interest and motivation, and external factors, such as teaching methods and school facilities (Astafani et al., 2024).

Green chemistry is a natural candidate for this kind of contextual teaching. It is defined as the design of chemical products and processes that reduce the use and generation of hazardous substances (Anastas & Eghbali, 2010). Its twelve principles, formulated by Anastas

and Warner (1998), direct learners to think about waste prevention, safer solvents, and renewable feedstocks rather than only about reaction procedures. Students nevertheless find the concepts difficult: a two-tier diagnostic survey of Chinese secondary students found their understanding of green chemistry concepts to be low (Ma & Hu, 2020). Teaching studies suggest that this can be improved. Green chemistry curricula produced conceptual gains among secondary school students in Malaysia (Karpudewan et al., 2015, 2016), and in Indonesia green chemistry has been brought into practicums, small-scale experiments, and basic chemistry courses in support of sustainability goals (Listyarini et al., 2019; Mitarlis et al., 2023; Redhana & Suardana, 2021). A quasi-experiment on reaction rate reported that a green chemistry practicum made an effective contribution of 28.3% to students' environmental awareness (Kamilah & Louise, 2025). This is a variance-based share from a posttest-only design, and it is not a standardised measure of effect size. Research on the field as a whole is still limited at the secondary level. In a systematic review of mainland Chinese research, Li and Eilks (2021) found that work on secondary education focused mainly on teaching strategies and that students' awareness of green chemistry was generally weak. Reviews of green chemistry education emphasise environmental consciousness and sustainable behaviour change as goals of teaching (Chen et al., 2020), and systematic reviews of its didactics continue to map the field (Sánchez Morales et al., 2024). Related work has examined growth in students' attitudes and understanding (Armstrong et al., 2018) and perspectives on education in green chemistry and sustainable chemistry (Zuin et al., 2021).

Ethnochemistry offers another route to relevance because it links the chemistry embedded in local cultural practice to school content. A literature review of Indonesian studies reported that the ethnochemistry approach was considered motivating for learning chemistry (Aldiansyah et al., 2023). An SDG-oriented ethno-STEM approach has been used to foster environmental awareness in elementary education (Pada et al., 2025), and an ethno-STEAM green chemistry e-module has been developed to improve senior high school students' environmental conservation skills (Epinur et al., 2026). Ethnochemistry-based culturally responsive teaching has been reported to improve cognitive outcomes in green chemistry at the senior high school level (Wardani et al., 2023). Culturally responsive teaching more generally treats students' cultural references as a resource for learning (Gay, 2002), and Mashami et al. (2025) explored the combination of green chemistry and cultural wisdom as a pathway to scientific literacy.

In this study the local practice was an infusion of Dayak onion (*Eleutherine palmifolia*) and ginger (*Zingiber officinale*), selected for its contextual relevance to students in Central Kalimantan. Dayak onion is an indigenous plant of Kalimantan that has been used for generations as folk medicine in Dayak communities (Harlita et al., 2018). An ethnobotanical survey of the Dayak Ot Danum and Tomum in Central Kalimantan lists ginger (*Z. officinale*, locally *jahe*) among the medicinal plants these communities use, and the ginger family (Zingiberaceae) was the family most often documented there (Neneng et al., 2025). In the lessons, students audited the preparation of the infusion against the twelve principles of green chemistry (Anastas & Warner, 1998).

What remains unclear is how a locally relevant context should be organised within the phases of Deep Learning so that students can work with green chemistry principles. Our rationale rests on three assumptions. First, a concrete local preparation gives students something to attach the principles to, which may make abstract ideas easier to hold; context-based chemistry education is built on this premise (Gilbert, 2006). Second, the meaningful phase asks students to judge the process against the principles, including its weak points, rather than to accept a list of principles. Third, the joyful phase closes with discussion and written reflection that link the findings to local environmental problems, a step that is identified as critical in Deep Learning in chemistry (Fatihah & Wiji, 2026). To the best of our knowledge and based on a review of recent literature, empirical studies that combine Deep Learning,

Dayak onion–ginger ethnochemistry, and simultaneous measurement of environmental awareness and green chemistry understanding remain scarce.

The study describes the baseline environmental awareness and understanding of green chemistry concepts of tenth-grade students at SMAN 1 Sampit, and compares the pretest–posttest gains in two intact classes, one taught with Dayak onion–ginger Deep Learning lessons and one taught conventionally. Specifically, this inquiry addresses two research questions: (1) What are the baseline levels of students' environmental awareness and green chemistry conceptual understanding before the intervention in both groups? and (2) Is there a noticeable difference in the normalised gain (N-gain) of environmental awareness and green chemistry understanding between students exposed to the ethnochemistry-based Deep Learning approach and those receiving conventional instruction? The contribution of the study is a descriptive account of score change in two intact classes that can inform the design of larger studies.

METHOD

Research Subjects

This study used a quantitative quasi-experimental pretest–posttest design with non-equivalent groups (Fraenkel et al., 2019). The target population was all tenth-grade students of SMAN 1 Sampit, and the accessible population was the students studying green chemistry in the odd semester. The unit of selection and of allocation was the intact class, and the unit of measurement was the student. Two tenth-grade classes with comparable prior chemistry achievement were selected purposively from the available classes on the basis of their mean scores on the chemistry daily assessment for the preceding instructional period (Experimental: Mean = 78.45, SD = 3.12; Control: Mean = 77.90, SD = 3.45). A simple lottery, performed once, then assigned one class to the Deep Learning condition (experimental class) and the other to conventional instruction on the same topic (control class). Class selection was therefore purposive, and only the allocation of the two selected classes to conditions was by lottery; students were not assigned individually. Each condition consisted of one class of 36 students, so 72 students were enrolled (Table 1). Both classes were taught by the same teacher in the same period. Item order was shuffled between pretest and posttest to reduce testing effects, and the classes were compared on the pretest before the intervention. The study took place at SMAN 1 Sampit, Kotawaringin Timur Regency, Central Kalimantan.

Table 1. Participant flow

Stage	Experimental class	Control class
Assigned to condition by lottery (class level)	1 class	1 class
Students enrolled in the class	36	36
Completed the pretest	36	36
Completed the posttest	36	36
Excluded or missing, with reasons	0	0
Participants with complete pretest and posttest data	36	36

Research Instruments

Environmental awareness was operationalised as an ecological worldview measured with the New Ecological Paradigm (NEP) scale (Dunlap & Van Liere, 1978; Dunlap, 2008). The questionnaire was adapted from the revised NEP scale (Dunlap et al., 2000) and contextualised to Central Kalimantan issues such as peatland fires, deforestation, river pollution, and the loss of endemic Bornean biodiversity. It contained 15 statements, three for each of five dimensions: reality of limits to growth, anti-anthropocentrism, fragility of nature's balance, rejection of exemptionalism, and possibility of an eco-crisis. Responses used a five-point Likert scale (1 = strongly disagree, 5 = strongly agree) with a mix of favourable and unfavourable statements, so total scores ranged from 15 to 75. The NEP scale measures an ecological worldview through self-report. In this article, environmental awareness refers to the NEP total score; it does not refer to observed behaviour. Dunlap et al. (2000) found that the 15 items could be treated as

one internally consistent summated scale, so only the total score is interpreted here. Specifically, eight statements (Items 1, 4, 7, 8, 10, 11, 13, and 14) were favourably worded, whereas seven statements (Items 2, 3, 5, 6, 9, 12, and 15) were unfavourably worded. Responses to the seven unfavourable items were reverse-coded (1 = 5, 2 = 4, 3 = 3, 4 = 2, 5 = 1) prior to summation so that higher total scores consistently reflected a stronger ecological worldview.

Understanding of green chemistry concepts was measured with a 20-item two-tier multiple-choice test, in which the first tier asks for a conceptual answer and the second tier for the reason (Treagust, 1988; Chandrasegaran et al., 2007). Items asked students to identify, analyse, and evaluate applications of the twelve green chemistry principles (Anastas & Warner, 1998) in the context of the Dayak onion–ginger infusion and in everyday situations relevant to Central Kalimantan. All 20 items were administered in the pretest and the posttest, and all 20 were scored. Each item was scored with a two-tier rule: 1 point was awarded only when both tiers were answered correctly, and 0 points otherwise. Total raw scores range from 0 to 20 and are reported on that scale; percentages (raw score ÷ 20 × 100) are given only where they help interpretation.

Instrument quality. Content validity of both instruments was judged by two experts (a lecturer in education and a senior chemistry teacher) using Aiken's V (Aiken, 1985). With two raters, $V = \Sigma s / [n(c - 1)]$, where n = 2 raters, c = number of rating categories, and s = the rating minus the lowest category. The experts used a four-category rating scale (c = 4) for both the NEP questionnaire and the concept test. With these parameters V can take only three values above 0.5: 0.667 (for example, both experts rated the item 3), 0.833 (ratings of 4 and 3), and 1.000 (both experts rated the item 4). Because only two experts rated the items, V is read here as a descriptive index of agreement and is not tested for significance. With only three discrete values of V above 0.5 possible (0.667, 0.833, 1.000), the threshold was set at $V \geq 0.667$ after expert discussion; items below this value were revised. After revision, Aiken's V for the concept test ranged from 0.667 to 0.833 (mean 0.767). For the NEP questionnaire, Aiken's V ranged from 0.833 to 1.000 (mean 0.978). Content-validity evidence was based on ratings from two experts.

Empirical validity was evaluated through try-outs involving 36 students per session, all of whom were distinct from the research sample. The first try-out of the concept test used students from a parallel tenth-grade cohort at the same school who had not yet studied the green chemistry unit; the second try-out (and the single NEP try-out) used a separate Class XI cohort. Item validity was determined using the corrected item–total correlation (r_{it}), obtained from the Corrected Item-Total Correlation output in SPSS. These values were compared against the critical Pearson correlation threshold ($r_{table} = 0.329$, two-tailed, $\alpha = 0.05$, df = 34). For the NEP questionnaire, item-total correlations were calculated after reverse-coding the seven unfavorable items. Reliability for both instruments was estimated using Cronbach's alpha, which is mathematically equivalent to KR-20 for dichotomously scored items (Tavakol & Dennick, 2011).

The green chemistry concept test underwent two trial rounds. In the initial try-out, 11 of the 20 items met the empirical validity criterion ($r_{it} \geq 0.329$). The remaining nine non-valid items (Items 1, 2, 4, 10, 11, 16, 17, 18, and 19) were revised based on item analysis and re-evaluated by the same two expert raters; details regarding the first try-out scores and specific item revisions are compiled in Table A4. A subsequent try-out with a separate Class XI cohort (n = 36), using the revised test, demonstrated that all 20 items satisfied the validity criterion ($r_{it} \geq 0.329$; Table A1), yielding a high internal consistency ($\alpha = 0.831$). The NEP questionnaire was administered once alongside the second concept test try-out (Table A3). All 15 NEP items met the empirical validity criterion ($r_{it} \geq 0.329$; Table A1), achieving strong reliability ($\alpha = 0.941$; Table A2). Detailed item-level statistics are provided in Appendix A.

Implementation was monitored with an observation sheet of six aspects, two for each phase (mindful, meaningful, joyful), each scored 1–4. The percentage score was interpreted as very good (86.00–100%), good (71.00–85.99%), fair (56.00–70.99%), or poor (<56.00%). The treatment was considered to have been delivered as designed when implementation reached at least 80%. The experimental class was observed with this sheet during the three Deep Learning meetings. The control class was observed using a separate conventional instruction sheet consisting of six parallel aspects structured around conventional teaching syntax (direct lesson orientation, prior knowledge activation, expository concept delivery, procedural worksheet practice, whole-class discussion, and teacher-led summarization), each rated on a 1–4 scale. Two independent observers who were blind to the research hypotheses rated both classes. Inter-rater agreement and implementation results are reported in the Results section.

Procedure

The experimental class attended three Deep Learning meetings (2 × 45 minutes per meeting), followed by the posttest in a fourth session that was not observed. The lessons followed the three Deep Learning phases in sequence and were built around preparing the Dayak onion–ginger infusion as a real context for green chemistry (Table 2). The control class studied the same green chemistry content through conventional instruction over the same number of meetings and duration (three meetings of 2 × 45 minutes each), in which the teacher led the lesson through explanation and procedural exercises. The pretest was given to both classes before the first meeting, and the posttest, using the same instruments, after the last meeting.

Table 2. Syntax of Deep Learning based on Dayak onion–ginger ethnochemistry

Phase	Student activities	Green chemistry connection
Mindful	Observe Dayak onion and ginger, identify their chemical constituents, and formulate a question about which green chemistry principles appear in preparing the infusion.	Activates prior knowledge of natural materials, safe solvents, and renewable feedstocks.
Meaningful	Prepare the infusion step by step, record and analyse each stage against the twelve principles, and evaluate the "greenness" of the process.	Direct identification of Principle 1 (waste prevention), Principle 5 (safer solvents; water), and Principle 7 (renewable feedstocks).
Joyful	Present and discuss the analyses, link the principles to environmental conservation in Central Kalimantan, and write a reflection.	Reinforces green chemistry as a bridge between science, local wisdom, and sustainability.

Data Analysis

Descriptive statistics (mean, standard deviation, minimum, and maximum) were computed for pretest and posttest scores in each class. Gain was taken as the posttest mean minus the pretest mean. N-Gain, (posttest − pretest) / (maximum score − pretest), with a maximum of 75 for the NEP and 20 for the concept test, was calculated for each student. Each student was classified into Hake's (1998) categories (low, $g < 0.30$; medium, $0.30 \leq g < 0.70$; high, $g \geq 0.70$), and the class-level N-gain was summarised as the mean of the individual values. Dimension (subscale) scores of the NEP were not analysed, because the scale is used here as a single summated scale (Dunlap et al., 2000). The analysis is descriptive. No inferential test of a treatment effect was applied because each condition included only one intact class.

RESULTS AND DISCUSSION

Implementation and Initial Condition

Implementation fidelity was 84.00% in the experimental class, above the 80% criterion. It was derived from meeting-level implementation scores of 81%, 84%, and 87% for Meetings 1, 2, and 3, respectively. Two independent observers blind to the hypotheses rated the experimental sessions with high agreement (Cohen's $\kappa = 0.88$ for the experimental class). The

same observers rated conventional instruction in the control class using a separate observation sheet, which yielded an implementation score of 70.13%. The reported κ = 0.88 refers only to the experimental class. Because the two sheets measure different aspects of teaching, these values are reported descriptively and are not compared.

Research Question 1. Before the intervention, the mean NEP score of all students was 49.06 out of 75 (65.41% of the maximum) and the mean concept score was 5.71 out of 20 (28.55%). The classes differed by 0.45 points on the NEP pretest (48.83 vs. 49.28), in favour of the control class, and by 0.08 points on the concept pretest (5.75 vs. 5.67), in favour of the experimental class (Table 3). These baseline differences were small.

Gains in Environmental Awareness and Concept Understanding

Table 3. Descriptive statistics of pretest, posttest, gain, and N-Gain by class (n = 36 per class)

Variable	Measure	Exp. M (SD)	Range	Con. M (SD)	Range
Environmental awareness (NEP, max 75)	Pretest	48.83 (8.08)	34–66	49.28 (6.55)	37–64
	Posttest	64.56 (5.80)	38–70	54.00 (7.36)	39–68
	Gain	15.72 (7.42)	−1–28	4.72 (4.83)	−5–17
	N-Gain	0.59 (0.18)	−0.03–0.82	0.19 (0.20)	−0.28–0.53
Green chemistry understanding (max 20)	Pretest	5.75 (4.07)	0–13	5.67 (4.01)	0–13
	Posttest	14.33 (5.12)	1–20	9.17 (5.79)	1–18
	Gain	8.58 (4.72)	−3–18	3.50 (3.93)	−6–13
	N-Gain	0.62 (0.31)	−0.23–1.00	0.27 (0.33)	−0.75–0.87

Note. n = 36 per class. Values are mean (standard deviation) and range (minimum–maximum), computed from individual student scores. Gain = posttest − pretest. N-Gain was calculated for each student as gain/(maximum score − pretest), with maximum scores of 75 (NEP) and 20 (concept test), and then averaged. NEP = New Ecological Paradigm.

Both classes scored higher on the posttest than on the pretest (Table 3 and Figure 1). The experimental class gained an average of 15.72 NEP points against 4.72 in the control class, and 8.58 concept-test points (out of 20) against 3.50. Averaged across students, N-Gain was 0.59 (SD = 0.18) and 0.19 (SD = 0.20) for environmental awareness, and 0.62 (SD = 0.31) and 0.27 (SD = 0.33) for concept understanding. Applying Hake's (1998) thresholds to these class means, the experimental class falls in the medium category and the control class in the low category on both outcomes. The ranges of the two classes overlap on every pretest and posttest variable (Table 3).

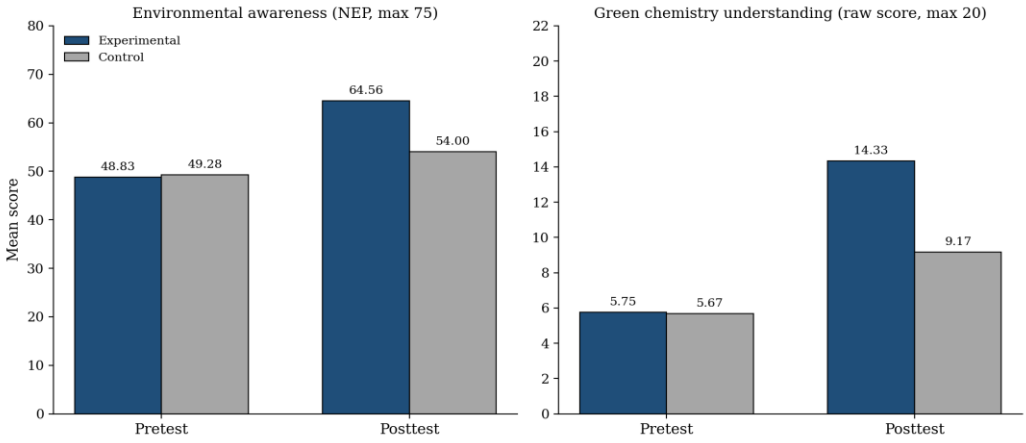


Figure 1. Mean pretest and posttest scores of the experimental and control classes (descriptive; n = 36 per class; concept test as raw score out of 20)

Individual N-Gain Categories

Class means hide how much students differed from one another. Table 4 shows the number and percentage of students in each of Hake's (1998) N-Gain categories, for each class and each variable, calculated from the individual student scores (n = 36 per class).

Table 4. N-Gain categories of individual students by class (n = 36 per class)

Hake category	Environmental awareness:		Concept understanding	
	Exp. n (%)	Control n (%)	Exp. n (%)	Control n (%)
Low ($g < 0.30$)	3 (8.3%)	25 (69.4%)	6 (16.7%)	20 (55.6%)
Medium ($0.30 \leq g < 0.70$)	22 (61.1%)	11 (30.6%)	11 (30.6%)	12 (33.3%)
High ($g \geq 0.70$)	11 (30.6%)	0 (0.0%)	19 (52.8%)	4 (11.1%)
Total	36 (100)	36 (100)	36 (100)	36 (100)

Note. Frequencies were counted from individual N-Gain values. Percentages are of n = 36 per class

In the experimental class, most students were in the medium or high category on both outcomes: 33 of 36 students (91.7%) for awareness and 30 (83.3%) for concept understanding. In the control class, most students were in the low category: 25 (69.4%) for awareness and 20 (55.6%) for concept understanding. Four control students (11.1%) reached a high gain on the concept test. The distributions overlap, because some experimental students gained little. These patterns describe the two classes.

Interpretation of the Findings

In these two classes, the mean gain in environmental awareness and in concept understanding was larger in the class taught with the Dayak onion–ginger Deep Learning lessons than in the class taught conventionally, and the same pattern appeared across individual N-gain categories (Table 4). A reading that fits both outcomes starts from the distinction between deep and surface approaches to learning, in which deep approaches aim at meaning and surface approaches at reproduction (Marton & Säljö, 1976); the Indonesian Deep Learning framework puts this aim into practice through mindful, meaningful, and joyful learning experiences (Kemendikdasmen, 2025). Each phase of the lessons can be related to the two outcomes in a different way. The mindful phase asks students to bring what they already know about natural materials to the infusion, which is consistent with the view that new understanding is built on prior knowledge (Bransford et al., 2000) and is the phase that may be most relevant for concept understanding. The meaningful phase requires students to audit each step of the preparation against the twelve principles, so the principles are used to judge a concrete process and are not memorised as a list; this is the logic of context-based chemistry teaching, which aims to make abstract chemistry more relevant (Gilbert, 2006), and it is the most plausible link to a two-tier test that asks for both an answer and its reason. The joyful phase connects the analyses to peatland fires, river pollution, and other local problems through discussion and reflection, which resembles socio-scientific reasoning about socially and environmentally consequential issues (Sadler, 2004) and is the phase closest to the ecological worldview that the NEP scale describes. That both scores moved is therefore compatible with the lesson design, but the data do not show which phase, if any, contributed, because the phases were neither varied nor measured separately; the limits on this inference are set out under Limitations.

Comparison with Previous Studies

The direction of the differences is consistent with earlier reports that green chemistry curricula improved secondary students' conceptual understanding (Karpudewan et al., 2015, 2016), that a green chemistry practicum was associated with higher environmental awareness (Kamilah & Louise, 2025), and that an ethnochemistry-based culturally responsive approach improved cognitive outcomes in green chemistry (Wardani et al., 2023). Score magnitudes are not compared with these studies because their designs and instruments differ. Karpudewan et al. (2015, 2016) evaluated green chemistry curricula and activities in Malaysian secondary schools. The present study differs in using a locally situated preparation, in adding an awareness outcome, and in analysing gains descriptively in two classes. Kamilah and Louise (2025) used a posttest-only quasi-experiment with cluster random sampling and reported a variance-based contribution rather than a standardised effect, and Wardani et al. (2023) measured cognitive outcomes with a posttest. Aldiansyah et al. (2023), a narrative review of

Indonesian studies that covered cultures from Java, Nusa Tenggara, Bali, Kalimantan, and Sumatra, concluded that the ethnochemistry approach is effective in motivating students to learn chemistry. That conclusion concerns motivation, which the present study did not measure, and a narrative review cannot estimate effects on understanding or awareness, so it supports the rationale for using a cultural context rather than the outcomes reported here. Pada et al. (2025) reported environmental awareness outcomes of an SDG-oriented ethno-STEM approach at the elementary level, and Epinur et al. (2026) developed an ethno-STEAM green chemistry e-module aimed at senior high school students' environmental conservation skills. Their value for the present study lies in the shared rationale of connecting local practice with environmental content, not in a comparison of effects.

Limitations

The main limitation is the design. Each condition consisted of one intact class, so the lessons, the class, the interaction between the teacher and each class, and differences in how fully each class was taught as planned cannot be separated, and the differences reported here are descriptive only. Student-level significance tests would treat 72 students as independent units although the condition was assigned at class level, and would therefore overstate precision; a small pretest difference does not show that the classes were equivalent in characteristics that were not measured. Other design limits are the single school, the short treatment (three meetings followed by a posttest session), and the absence of a delayed posttest, so retention could not be examined. Implementation was observed with different sheets in the two classes, so fidelity could not be compared. The study also had no culturally neutral green chemistry comparison and therefore gives no evidence that these lessons work better than green chemistry teaching without a local context.

Two further points concern the local context. Contextual relevance was inferred from the ethnobotanical literature and was not measured as a separate variable. The cited sources document the medicinal use of each plant in Dayak communities (Harlita et al., 2018; Neneng et al., 2025), not the combined infusion as a school context. Moreover, Neneng et al. (2025) surveyed Dayak Ot Danum and Tomum communities in Katingan and Lamandau districts rather than Kotawaringin Timur (Sampit); cultural relevance for students at SMAN 1 Sampit is therefore indirect. In addition, a herbal preparation is not green because its ingredients are natural; whether it is green depends on how it is made. Water as the solvent and use of the whole bulb point towards safer solvents and waste prevention, whereas the heating step consumes energy and the plants may or may not come from renewable sources (Anastas & Warner, 1998). The infusion was therefore used as a case to be audited against the principles and not as an example that is green by definition.

The concept-test scores need careful reading. The test is researcher-made, and its validity evidence is limited to expert judgement and two try-outs. Nine of its 20 items did not meet the validity criterion in the first try-out. They were revised and tried out again before use. The items also reused the infusion context that the experimental class had worked with, so a higher score may reflect near transfer of practised content rather than far transfer of green chemistry principles to unfamiliar processes. Barnett and Ceci (2002) distinguish near from far transfer by how far learning is applied across contexts. The two-tier format, which requires both a correct answer and a correct reason (Chandrasegaran et al., 2007; Treagust, 1988), reduces guessing but does not address this issue. Separating the explanations would require items set in contexts that neither class studied, scored blind to condition.

The NEP scale captures an ecological worldview through self-report (Dunlap et al., 2000; Dunlap, 2008). A higher score shows stronger endorsement of the statements and may be affected by social desirability. The score therefore describes change in ecological worldview and not change in behaviour. Gaps between environmental attitudes and behaviour are well documented (Kollmuss & Agyeman, 2002), and models of responsible environmental behaviour treat awareness-level variables as only one group among several, alongside

ownership and empowerment variables (Hungerford & Volk, 1990). This article therefore makes no claim about behavioural outcomes. Gain-based indices depend on the pretest score, so the N-gain values of the two classes should be read with that in mind. Marx and Cummings (2007) discuss normalised change as an alternative index for this reason. In this article, Hake's (1998) categories are used descriptively.

CONCLUSION

This preliminary study described the environmental awareness and green chemistry understanding of two tenth-grade classes at SMAN 1 Sampit. Before the intervention, the mean NEP score was 49.06 out of 75 and the mean concept score was 5.71 out of 20. Over the three Deep Learning meetings, the class taught with the Dayak onion–ginger lessons showed a larger mean individual N-gain than the conventionally taught class on both measures (0.59 versus 0.19 for ecological worldview; 0.62 versus 0.27 for concept understanding). These descriptive observations document the pattern of score change in two intact classes and can inform the design of larger studies. They cannot be interpreted as estimates of a causal effect or as evidence of transfer to unfamiliar chemical processes.

RECOMMENDATION

Teachers who want to try this kind of lesson can use the infusion as a case that students audit against the green chemistry principles, including its weak points such as energy use and the source of the plants, and should keep laboratory safety in view. Future studies should assign several classes per condition at random and analyse outcomes at the class level or with multilevel models, which also requires effect sizes that account for clustering (Hedges, 2007). Effect sizes should be reported with their confidence intervals (Lakens, 2013; Tomczak & Tomczak, 2014). Studies should include a culturally neutral green chemistry comparison, observe both conditions with the same instrument, use a fully validated concept test with items from unfamiliar contexts and blind scoring, add a delayed posttest and behavioural indicators, and document parental consent, student assent, and ethics-committee review or exemption in accordance with applicable requirements. Anonymised paired scores should be made available for audit.

ACKNOWLEDGMENT

The authors thank the students and teachers of SMAN 1 Sampit, the two expert validators, and the supervisors at Universitas Muhammadiyah Malang for their support.

FUNDING INFORMATION

This research did not receive external funding.

AUTHOR CONTRIBUTIONS STATEMENT

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Anna Farida	✓	✓			✓	✓	✓	✓	✓	✓	✓		✓	
Dyah W. Ekowati	✓	✓				✓		✓		✓	✓	✓		
Agus Tinus	✓		✓							✓		✓		

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

ETHICAL APPROVAL

The study was conducted with the written permission of the principal of SMAN 1 Sampit via the issued research permit (No. 400.3.5/205/14/SMAN 1 spt/VIII/2026). In addition, it was supported by a research permit letter (No. E.5.d/764/DPPs-UMM/VII/2026) issued by the Directorate of Postgraduate Programs, Universitas Muhammadiyah Malang, and addressed to the school. Students and their parents or guardians were informed about the purpose and procedures of the research and gave written informed consent, and students participated voluntarily in the tests and questionnaires. Student identities were anonymized using respondent codes (Exp-01 to Exp-37 for the experimental class and Con-01 to Con-37 for the control class), and the analysis files contain only these codes. No ethics-committee review was conducted for this study; research permission was obtained as described above. Anonymized item-level data (Supplementary Dataset S1) and the analysis script (Supplementary File S5) are provided as Supplementary Material.

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request. The data are not publicly available due to privacy considerations of research participants. Anonymised paired student scores can be made available to the editor and reviewers upon request for audit purposes.

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

Aldiansyah, A., Pasa, J. I., Muttaqin, M. R., Awaliyah, N. N., & Erika, F. (2023). Literatur review: Keterkaitan pembelajaran kimia terhadap pendekatan etnokimia di Indonesia [Literature review: The relationship between chemistry learning and the ethnochemistry approach in Indonesia]. *CHEDS: Journal of Chemistry, Education, and Science, 7*(2), 238–246. https://doi.org/10.30743/cheds.v7i2.8416

Anastas, P. T., & Eghbali, N. (2010). Green chemistry: Principles and practice. *Chemical Society Reviews, 39*(1), 301–312. https://doi.org/10.1039/B918763B

Anastas, P. T., & Warner, J. C. (1998). *Green chemistry: Theory and practice*. Oxford University Press.

Armstrong, L. B., Rivas, M. C., Douskey, M. C., & Baranger, A. M. (2018). Teaching students the complexity of green chemistry and assessing growth in attitudes and understanding. *Current Opinion in Green and Sustainable Chemistry, 13*, 61–67. https://doi.org/10.1016/j.cogsc.2018.04.003

Astafani, A., Resmawati, R. F., & Hakim, M. E. L. (2024). Systematic review: Faktor-faktor kesulitan belajar materi kimia [Systematic review: Factors contributing to difficulties in learning chemistry]. *Jurnal Inovasi Pendidikan Kimia, 18*(2), 81–88. https://journal.unnes.ac.id/journals/JIPK/article/view/5458

Barnett, S. M., & Ceci, S. J. (2002). When and where do we apply what we learn? A taxonomy for far transfer. *Psychological Bulletin, 128*(4), 612–637. https://doi.org/10.1037/0033-2909.128.4.612

Bransford, J. D., Brown, A. L., & Cocking, R. R. (Eds.). (2000). *How people learn: Brain, mind, experience, and school* (Expanded ed.). National Academy Press. https://doi.org/10.17226/9853

Chandrasegaran, A. L., Treagust, D. F., & Mocerino, M. (2007). The development of a two-tier multiple-choice diagnostic instrument for evaluating secondary school students' ability to describe and explain chemical reactions using multiple levels of representation. *Chemistry Education Research and Practice, 8*(3), 293–307. https://doi.org/10.1039/B7RP90006F

Chen, M., Jeronen, E., & Wang, A. (2020). What lies behind teaching and learning green chemistry to promote sustainability education? A literature review. *International Journal of Environmental Research and Public Health, 17*(21), Article 7876. https://doi.org/10.3390/ijerph17217876

Dunlap, R. E. (2008). The New Environmental Paradigm Scale: From marginality to worldwide use. *The Journal of Environmental Education, 40*(1), 3–18. https://doi.org/10.3200/JOEE.40.1.3-18

Dunlap, R. E., & Van Liere, K. D. (1978). The "new environmental paradigm." *The Journal of Environmental Education, 9*(4), 10–19.

Dunlap, R. E., Van Liere, K. D., Mertig, A. G., & Jones, R. E. (2000). Measuring endorsement of the New Ecological Paradigm: A revised NEP scale. *Journal of Social Issues, 56*(3), 425–442. https://doi.org/10.1111/0022-4537.00176

Epinur, Fuldiaratman, Nelyahardi, Hendra, & Minarni. (2026). Development of a green chemistry e-module integrated with ethno-STEAM for enhancing senior high school students' environmental conservation skills. *Jurnal Penelitian Pendidikan IPA, 12*(1),

513–520. https://doi.org/10.29303/jppipa.v12i1.13024

Fatihah, W., & Wiji, W. (2026). Implementasi deep learning pada pembelajaran kimia: Tantangan, urgensi dan transformasi pedagogik [Implementation of deep learning in chemistry learning: Challenges, urgency, and pedagogical transformation]. *Jurnal Riset dan Praktik Pendidikan Kimia, 14*(1). https://ejournal.upi.edu/index.php/JRPPK/article/view/96009

Fraenkel, J. R., Wallen, N. E., & Hyun, H. H. (2019). *How to design and evaluate research in education* (10th ed.). McGraw-Hill Education.

Fullan, M., Quinn, J., & McEachen, J. (2018). *Deep learning: Engage the world change the world*. Corwin.

Gay, G. (2002). Preparing for culturally responsive teaching. *Journal of Teacher Education, 53*(2), 106–116. https://doi.org/10.1177/0022487102053002003

Gilbert, J. K. (2006). On the nature of "context" in chemical education. *International Journal of Science Education, 28*(9), 957–976. https://doi.org/10.1080/09500690600702470

Hake, R. R. (1998). Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *American Journal of Physics, 66*(1), 64–74. https://doi.org/10.1119/1.18809

Harlita, T. D., Oedjijono, & Asnani, A. (2018). The antibacterial activity of Dayak onion (*Eleutherine palmifolia* (L.) Merr) towards pathogenic bacteria. *Tropical Life Sciences Research, 29*(2), 39–52. https://doi.org/10.21315/tlsr2018.29.2.4

Hedges, L. V. (2007). Effect sizes in cluster-randomized designs. *Journal of Educational and Behavioral Statistics, 32*(4), 341–370. https://doi.org/10.3102/1076998607302626

Hungerford, H. R., & Volk, T. L. (1990). Changing learner behavior through environmental education. *The Journal of Environmental Education, 21*(3), 8–21. https://doi.org/10.1080/00958964.1990.10753743

Kamilah, I., & Louise, I. S. Y. (2025). Pengaruh praktikum green chemistry terhadap keterampilan proses sains dan kesadaran lingkungan pada materi faktor laju reaksi [The effect of green chemistry practicums on science process skills and environmental awareness in reaction rate factors]. *Jurnal Pendidikan Matematika dan Sains, 13*(Special Issue), 247–259. https://doi.org/10.21831/jpms.v13iSpecial_issue.89688

Karpudewan, M., Roth, W.-M., & Ismail, Z. (2015). The effects of "green chemistry" on secondary school students' understanding and motivation. *The Asia-Pacific Education Researcher, 24*(1), 35–43. https://doi.org/10.1007/s40299-013-0156-z

Karpudewan, M., Roth, W.-M., & Sinniah, D. (2016). The role of green chemistry activities in fostering secondary school students' understanding of acid–base concepts and argumentation skills. *Chemistry Education Research and Practice, 17*(4), 893–901. https://doi.org/10.1039/c6rp00079g

Kementerian Pendidikan Dasar dan Menengah. (2025). *Naskah akademik pembelajaran mendalam: Menuju pendidikan bermutu untuk semua* [Academic paper on deep learning: Towards quality education for all]. Pusat Kurikulum dan Pembelajaran, Badan Standar, Kurikulum, dan Asesmen Pendidikan. https://statik.unesa.ac.id/profileunesa_konten_statik%2Fuploads%2Fs1pkimia/file/e760a054-f766-4f33-83b6-89e7c62719b2.pdf

Kollmuss, A., & Agyeman, J. (2002). Mind the gap: Why do people act environmentally and what are the barriers to pro-environmental behavior? *Environmental Education Research, 8*(3), 239–260. https://doi.org/10.1080/13504620220145401

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs. *Frontiers in Psychology, 4*, Article 863. https://doi.org/10.3389/fpsyg.2013.00863

Li, B., & Eilks, I. (2021). A systematic review of green and sustainable chemistry education research literature in mainland China. *Sustainable Chemistry and Pharmacy, 21*, Article

100446. https://doi.org/10.1016/j.scp.2021.100446

Listyarini, R. V., Pamenang, F. D. N., Harta, J., Wijayanti, L. W., Asy'ari, M., & Lee, W. (2019). The integration of green chemistry principles into small scale chemistry practicum for senior high school students. *Jurnal Pendidikan IPA Indonesia, 8*(3), 371–378. https://doi.org/10.15294/jpii.v8i3.19250

Ma, J., & Hu, S. (2020). Evaluating Chinese secondary school students' understanding of green chemistry. *Science Education International, 31*(2), 209–219. https://doi.org/10.33828/sei.v31.i2.11

Marton, F., & Säljö, R. (1976). On qualitative differences in learning: I—Outcome and process. *British Journal of Educational Psychology, 46*(1), 4–11. https://doi.org/10.1111/j.2044-8279.1976.tb02980.x

Marx, J. D., & Cummings, K. (2007). Normalized change. *American Journal of Physics, 75*(1), 87–91. https://doi.org/10.1119/1.2372470

Mashami, R. A., Ahmadi, & Pahriah. (2025). Green chemistry and cultural wisdom: A pathway to improving scientific literacy among high school students. *Social Sciences & Humanities Open, 11*, Article 101653. https://doi.org/10.1016/j.ssaho.2025.101653

Mitarlis, Azizah, U., & Yonata, B. (2023). The integration of green chemistry principles in basic chemistry learning to support achievement of sustainable development goals (SDGs) through education. *Journal of Technology and Science Education, 13*(1), 233–254. https://doi.org/10.3926/jotse.1892

Neneng, L., Nawan, Naibaho, F. G., Rahman, S., Septya, L., Panjaitan, D., Lestari, R. W., Irsyam, A. S. D., Norwahyunie, & Noraini. (2025). Botanical survey of ethnomedicines in Dayak sub-ethnic groups, Ot Danum and Tomum in Central Kalimantan, Indonesia. *Biodiversitas, 26*(1), 458–469. https://doi.org/10.13057/biodiv/d260145

OECD. (2023). *PISA 2022 results (Volume I): The state of learning and equity in education*. OECD Publishing. https://doi.org/10.1787/53f23881-en

Pada, A., Chanunan, S., & Rahmat, I. (2025). Fostering environmental awareness through sustainable development goal-oriented ethno-STEM approach in elementary education. *Jurnal Pendidikan IPA Indonesia, 14*(3), 469–479. https://doi.org/10.15294/jpii.v14i3.24420

Redhana, I. W., & Suardana, I. N. (2021). Green chemistry practicums at chemical equilibrium shift to enhance students' learning outcomes. *International Journal of Instruction, 14*(1), 691–708. https://doi.org/10.29333/IJI.2021.14142A

Sadler, T. D. (2004). Informal reasoning regarding socioscientific issues: A critical review of research. *Journal of Research in Science Teaching, 41*(5), 513–536. https://doi.org/10.1002/tea.20009

Sánchez Morales, R., Sáenz-López, P., & de las Heras Perez, M. A. (2024). Green chemistry and its impact on the transition towards sustainable chemistry: A systematic review. *Sustainability, 16*(15), Article 6526. https://doi.org/10.3390/su16156526

Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education, 2*, 53–55. https://doi.org/10.5116/ijme.4dfb.8dfd

Tomczak, M., & Tomczak, E. (2014). The need to report effect size estimates revisited: An overview of some recommended measures of effect size. *Trends in Sport Sciences, 21*(1), 19–25.

Treagust, D. F. (1988). Development and use of diagnostic tests to evaluate students' misconceptions in science. *International Journal of Science Education, 10*(2), 159–169. https://doi.org/10.1080/0950069880100204

Wardani, L. K., Mulyani, B., Ariani, S. R. D., & Yamtinah, S. (2023). The effect of an ethnochemistry-based culturally responsive teaching approach to improve cognitive learning outcomes on green chemistry material in high school. *Jurnal Penelitian Pendidikan IPA, 9*(12), 11029–11037. https://doi.org/10.29303/jppipa.v9i12.5532

Zuin, V. G., Eilks, I., Elschami, M., & Kümmerer, K. (2021). Education in green chemistry and in sustainable chemistry: Perspectives towards sustainability. *Green Chemistry, 23*(4), 1594–1608. https://doi.org/10.1039/d0gc03313h

APPENDIX

ITEM-LEVEL VALIDITY STATISTICS FROM THE TRY-OUTS (n = 36 PER TRY-OUT)

Table A1. Item validity results of the second try-out (n = 36): corrected item–total correlation (r_{it})

No	Item	r_{it}	r_{table}	Decision	Remarks
A. Chemistry test (two-tier, scored 0/1; 20 items)					
1	Q1	0.474	0.329	Valid	Retained
2	Q2	0.582	0.329	Valid	Retained
3	Q3	0.446	0.329	Valid	Retained
4	Q4	0.480	0.329	Valid	Retained
5	Q5	0.571	0.329	Valid	Retained
6	Q6	0.540	0.329	Valid	Retained
7	Q7	0.529	0.329	Valid	Retained
8	Q8	0.554	0.329	Valid	Retained
9	Q9	0.411	0.329	Valid	Retained
10	Q10	0.428	0.329	Valid	Retained
11	Q11	0.484	0.329	Valid	Retained
12	Q12	0.415	0.329	Valid	Retained
13	Q13	0.603	0.329	Valid	Retained
14	Q14	0.446	0.329	Valid	Retained
15	Q15	0.438	0.329	Valid	Retained
16	Q16	0.407	0.329	Valid	Retained
17	Q17	0.567	0.329	Valid	Retained
18	Q18	0.472	0.329	Valid	Retained
19	Q19	0.437	0.329	Valid	Retained
20	Q20	0.476	0.329	Valid	Retained
B. NEP Questionnaire (1–5 scale; 15 items)					
1	NEP Q1	0.661	0.329	Valid	Retained
2	NEP Q2	0.600	0.329	Valid	Retained
3	NEP Q3	0.545	0.329	Valid	Retained
4	NEP Q4	0.683	0.329	Valid	Retained
5	NEP Q5	0.663	0.329	Valid	Retained
6	NEP Q6	0.659	0.329	Valid	Retained
7	NEP Q7	0.725	0.329	Valid	Retained
8	NEP Q8	0.810	0.329	Valid	Retained
9	NEP Q9	0.687	0.329	Valid	Retained
10	NEP Q10	0.784	0.329	Valid	Retained
11	NEP Q11	0.785	0.329	Valid	Retained
12	NEP Q12	0.701	0.329	Valid	Retained
13	NEP Q13	0.656	0.329	Valid	Retained
14	NEP Q14	0.805	0.329	Valid	Retained
15	NEP Q15	0.697	0.329	Valid	Retained

Note. r_{it} = corrected item–total correlation, taken from the Corrected Item-Total Correlation column of the SPSS reliability output. Items were judged valid when $r_{it} \geq r_{table} = 0.329$ (two-tailed, $\alpha = 0.05$, df = 34, N = 36). Observed r_{it} ranged from 0.407 to 0.603 for the concept test and from 0.545 to 0.810 for the NEP questionnaire. For the NEP questionnaire, correlations were computed after reverse-coding the unfavourable items (2, 3, 5, 6, 9, 12, and 15).

Table A2. Summary of Instrument Reliability Test Results

Instrument	N	Number of Items	Cronbach's Alpha	Criterion	Remarks
Green Chemistry Test	36	20	0.831	High	Reliable
NEP Questionnaire	36	15	0.941	Very high	Reliable

Note: Reliability was estimated using Cronbach's alpha. For dichotomous tests (0/1), this coefficient is equivalent to KR-20. An instrument is considered reliable if $\alpha \geq 0.60$.

Table A3. Summary of Valid Items

Instrument	Trial	Number of Items	Valid	Invalid	Percentage Valid
Green Chemistry Test	First try-out	20	11	9	55%
Green Chemistry Test	Second try-out (revised)	20	20	0	100%
NEP Questionnaire	Single try-out	15	15	0	100%

Note: Item analysis was performed in SPSS (N = 36 respondents per try-out). The NEP questionnaire was tried out once, together with the second try-out of the concept test. Percentages refer to items within each instrument and trial.

Table A4. Item validity results of the first try-out of the concept test (n = 36): corrected item–total correlation (r_{it}), decision, and revision of the nine items that did not meet the criterion

No	r_{it}	r_{table}	Decision	Remarks
1	0.140	0.329	Invalid	Revised: Distractor clarity improved in Tier 1
2	0.314	0.329	Invalid	Revised: Simplified language phrasing in Tier 2
3	0.446	0.329	Valid	Retained
4	0.160	0.329	Invalid	Revised: Aligned item scenario with Principle 1
5	0.417	0.329	Valid	Retained
6	0.574	0.329	Valid	Retained
7	0.554	0.329	Valid	Retained
8	0.528	0.329	Valid	Retained
9	0.395	0.329	Valid	Retained
10	0.188	0.329	Invalid	Revised: Replaced ambiguous option choices
11	0.223	0.329	Invalid	Revised: Refined regional context relevance
12	0.394	0.329	Valid	Retained
13	0.479	0.329	Valid	Retained
14	0.501	0.329	Valid	Retained
15	0.387	0.329	Valid	Retained
16	0.223	0.329	Invalid	Revised: Clarified technical chemical terminology
17	-0.222	0.329	Invalid	Revised: Corrected logical link between tiers
18	0.132	0.329	Invalid	Revised: Adjusted difficulty level of prompt
19	0.295	0.329	Invalid	Revised: Rephrased prompt to eliminate vagueness
20	0.462	0.329	Valid	Retained

Note: Eleven items met the criterion (r_{it} = 0.387–0.574); nine items (1, 2, 4, 10, 11, 16, 17, 18, and 19) did not and were revised and re-evaluated by the same two experts before the second try-out (Table A1). Item analysis was performed in SPSS.