

Lpt UPGRIS

979 page 397-413

 LPT February 2026 V5I1

Document Details

Submission ID**trn:oid::3618:156066507****Submission Date****Oct 27, 2026, 1:38 PM GMT+7****Download Date****Oct 27, 2026, 1:44 PM GMT+7****File Name****979 page 397-413.pdf****File Size****585.8 KB****17 Pages****8,147 Words****49,852 Characters**

7% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.

Filtered from the Report

- Bibliography
- Quoted Text
- Small Matches (less than 8 words)

Match Groups

-  **26 Not Cited or Quoted 6%**
Matches with neither in-text citation nor quotation marks
-  **6 Missing Quotations 1%**
Matches that are still very similar to source material
-  **0 Missing Citation 0%**
Matches that have quotation marks, but no in-text citation
-  **0 Cited and Quoted 0%**
Matches with in-text citation present, but no quotation marks

Top Sources

- 6%  Internet sources
- 3%  Publications
- 2%  Submitted works (Student Papers)

Match Groups

- 26 Not Cited or Quoted 6%**
Matches with neither in-text citation nor quotation marks
- 6 Missing Quotations 1%**
Matches that are still very similar to source material
- 0 Missing Citation 0%**
Matches that have quotation marks, but no in-text citation
- 0 Cited and Quoted 0%**
Matches with in-text citation present, but no quotation marks

Top Sources

- 6% Internet sources
- 3% Publications
- 2% Submitted works (Student Papers)

Top Sources

The sources with the highest number of matches within the submission. Overlapping sources will not be displayed.

1	Internet	journal3.upgris.ac.id	<1%
2	Internet	indochembull.com	<1%
3	Internet	www.researchsquare.com	<1%
4	Internet	journal.upgris.ac.id	<1%
5	Internet	paeale.conferenceservice.jp	<1%
6	Internet	www.ojs.sijm.it	<1%
7	Internet	mdpi-res.com	<1%
8	Internet	arxiv.org	<1%
9	Internet	jele.or.id	<1%
10	Publication	Risky Fitri Anggraini, Lisanul Uswah Sadieda, Nurul Hidayati. "PENGEMBANGAN ...	<1%

11	Internet	experientiallearningedac634.blogspot.com	<1%
12	Internet	inkwellpub.com	<1%
13	Internet	journal.unismuh.ac.id	<1%
14	Internet	jppipa.unram.ac.id	<1%
15	Internet	jurnal.unissula.ac.id	<1%
16	Internet	download.atlantis-press.com	<1%
17	Internet	journal-fip.um.ac.id	<1%
18	Internet	pmc.ncbi.nlm.nih.gov	<1%
19	Internet	www.researchgate.net	<1%
20	Student papers	Mahidol University on 2009-12-14	<1%
21	Internet	academicleadership.org	<1%
22	Internet	carrels.distantreader.org	<1%
23	Internet	e-journal.unair.ac.id	<1%
24	Internet	journal.ia-education.com	<1%

25 Internet

journals-jd.upm.edu.my

<1%

26 Internet

www.mdpi.com

<1%

Implementing a Five-Phase Experiential-Based Deep Learning Model to Support Students' Conceptual Understanding of Force in Secondary School

Christo Vincentius Goni¹, Orbanus Naharia², Marianus Marianus³, Kenny Setiawan Lahope^{3*}, Tineke Makahinda³ and Darius Ladbam Lihu⁴

¹ Science Education Program, Postgraduate Program, Universitas Negeri Manado, Jl. Kampus Unima, Tonsaru, Minahasa, North Sulawesi, Indonesia.

² Department of Biology, Universitas Negeri Manado, Jl. Kampus Unima, Tonsaru, Minahasa, North Sulawesi, Indonesia.

³ Department of Physics, Universitas Negeri Manado, Jl. Kampus Unima, Tonsaru, Minahasa, North Sulawesi, Indonesia.

⁴ Department of Electrical Engineering Education Program, Universitas Negeri Manado, Jl. Kampus Unima, Tonsaru, Minahasa, North Sulawesi, Indonesia.

*Corresponding author(s), e-mail: kennylahope15@gmail.com

Article Info:

Article History:

received: 14 July 2026

accepted: 20 October 2026

available online: 27 October 2026

Keyword:

Experiential-Based Deep Learning (EBDL), conceptual understanding, force, secondary school physics, experiential learning, reflective inquiry

<https://doi.org/10.26877/lpt.v5i3.979>

Copyright © 2026 by author(s).

Published by Lontar Physics Today. This is an open access article under the CC BY license (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



OPEN ACCESS

Abstract:

Students' conceptual understanding of force is challenged by persistent difficulties in interpreting interactions, resultant forces, and changes in motion. This study examined the classroom implementation of a five-phase Experiential-Based Deep Learning (EBDL) instructional model designed to connect firsthand observation with reflection, evidence-based explanation, and transfer. A quasi-experimental nonequivalent pretest–posttest design involved 40 eighth-grade students at one Indonesian junior secondary school, with 20 students receiving EBDL and 20 receiving Direct Instruction. Conceptual understanding was assessed using five essay items on force; expert content ratings yielded an overall Aiken's V of 0.870 after recalculation. Available group-level posttest results favored the EBDL group ($M = 85.85$, $SD = 5.51$) over the control group ($M = 81.30$, $SD = 7.00$). The reported Mann–Whitney test indicated a between-group difference ($U = 124.00$, $Z = -2.06$, $p = 0.039$, $r = 0.326$). The model's distinctive instructional contribution is its explicit progression from discrepant experiences through reflective meaning-making and force-based reasoning to contextual application and transfer. However, aggregate posttest results do not establish which cognitive process benefited most, whether individual misconceptions were corrected, or the magnitude of pretest-to-posttest change. These findings support further testing of EBDL as a structured, context-sensitive alternative for junior secondary force instruction, with stronger item-level and implementation evidence required before wider causal or generalizable claims are made.

1. Introduction

Conceptual understanding is central to physics education because learners must use physical principles to interpret phenomena, justify predictions, and apply relationships beyond familiar exercises. This expectation is reflected in scientific literacy frameworks that emphasize explaining phenomena and interpreting evidence (OECD, 2023; Docktor & Mestre, 2014). In mechanics, understanding force requires more than selecting equations: students must distinguish interactions from motion, determine the resultant force, and reason about acceleration and equilibrium (Hestenes et al., 1992; Savinainen & Scott, 2002).

Nevertheless, students frequently retain intuitive ideas about force after instruction. Familiar misconceptions include treating motion as evidence of a continuously acting force and assuming that equal and opposite forces must always stop an object. Such reasoning conflicts with Newtonian mechanics, in which acceleration depends on the net external force, whereas zero net force is compatible with rest or constant velocity (Hestenes et al., 1992; Chrysostomou et al., 2024). Physics education

research therefore identifies conceptual interpretation—not simply procedural attainment—as a critical instructional target (Docktor & Mestre, 2014).

Expository teaching and worked examples can efficiently introduce physics terminology and procedures, but they do not by themselves ensure that students compare prior predictions with physical evidence or explain why a mathematical result is reasonable. Studies of interactive engagement in mechanics show the importance of opportunities to articulate and test reasoning (Hake, 1998; Savinainen & Scott, 2002). This study does not assume that Direct Instruction is inherently ineffective; rather, it asks whether a structured sequence of experience, explanation, and transfer is associated with stronger force-concept performance under the studied classroom conditions.

Force provides a suitable domain for this inquiry because junior secondary students encounter pushes and pulls daily yet may struggle to connect these experiences with resultant force, Newton's second law, and balanced forces. The instrument in Appendix A targets opposing forces on a table, a kicked ball, the acceleration of a cart, balanced pulls on a box, and collaborative pushing of a wardrobe. These tasks invite conceptual reasoning about the sign and direction of forces, changes in motion, equilibrium, and the dependence of acceleration on net force.

Addressing such difficulties requires instructional interactions in which students make predictions, observe the consequences of an applied force, compare competing explanations, and reconstruct conclusions using physical evidence. Direct experience alone is not sufficient: teachers must also help learners distinguish the observation from its scientifically valid interpretation (Morris, 2020; Docktor & Mestre, 2014).

Preliminary classroom observations at SMP Negeri 1 Tengah, as described in the initial field account, indicated reliance on explanation and formula practice and limited opportunities for practical investigation and reflection. These observations informed the local instructional problem; they were not a systematically analyzed baseline dataset. The relevant research problem is therefore whether a force-focused, experience-to-reasoning teaching sequence can support conceptual understanding in this setting.

Kolb's experiential learning theory distinguishes four foundational processes—concrete experience, reflective observation, abstract conceptualization, and active experimentation (Kolb, 2015). Those processes inform, but are not identical to, the instructional model tested here. EBDL has five pedagogical phases: Mindful Experiencing, Reflective Meaning-Making, Deep Conceptualization, Joyful Knowledge Application, and Transformative Reflection and Transfer. The final phase explicitly revisits and extends explanations in unfamiliar contexts.

In EBDL, deep learning refers to meaningful intellectual engagement in the education sense, not to an artificial-intelligence technique. The model combines careful attention to a phenomenon, evidence-based discussion, explicit concept building, contextual problem solving, and transfer. The pedagogical hypothesis is that the transition from an intuitive force explanation to a reasoned account of net force is supported by the sequence as a whole; the present dataset cannot isolate the causal contribution of any individual phase (Fullan et al., 2018; Docktor & Mestre, 2014).

Although experiential and interactive approaches have been investigated in physics education, a practical problem for teachers is translating these principles into an explicit routine linked to particular force concepts and cognitive demands. Research on interactive engagement and conceptual inventories motivates the use of prediction, empirical checks, and explanation, while EBDL specifies when each is used and how a final transfer task closes the cycle (Hake, 1998; Savinainen & Scott, 2002).

The study addresses a local instructional and operational gap: limited force-specific guidance for moving systematically from students' everyday interpretations to evidence-based reasoning about resultant forces and motion. Its proposed contribution is a clearly described instructional sequence that educators can adapt and evaluate. It does not claim that the underlying ideas of experience, reflection, or conceptual change were previously absent from physics education research.

The EBDL design is derived from, but does not replace, Kolb's four-part theoretical cycle. The five EBDL phases are classroom enactments with identifiable teacher prompts, student reasoning tasks, instructional supports, and intended cognitive outcomes. In contrast, the control lessons followed the stated Direct Instruction sequence of objectives, teacher explanation, guided practice, independent practice, and evaluation. The comparison thus concerns two documented instructional organizations applied to the same topic, not an opposition between learning and non-learning.

The potentially novel contribution is the force-specific operationalization of these principles: students first make and check predictions about pushes and pulls, then distinguish observed motion from net force, formalize the vector relationship between forces and acceleration, justify competing solutions, and transfer that reasoning to an unfamiliar context. This sequence offers a testable instructional design; its individual mechanisms remain hypotheses until process data and item-level responses are analyzed.

Accordingly, this study aimed (a) to specify the rationale, phases, support system, and enactment of the EBDL model for the junior secondary topic of force and (b) to examine whether students taught with EBDL exhibited higher post-instruction conceptual-understanding scores than a locally comparable Direct Instruction group. The study also discusses the educational plausibility, boundaries, and requirements for more rigorous evaluation of the model.

2. Theoretical Framework

2.1. *Constructivist Principles for Force Understanding*

Students interpret pushes, pulls, and moving objects using prior experiences that may not correspond to Newtonian principles. From a constructivist perspective, learning requires opportunities to make these intuitive explanations explicit and evaluate them against evidence rather than merely hear correct rules (Piaget, 1977; Docktor & Mestre, 2014). In EBDL, prediction and explanation precede formalization so that the teacher can identify which force relationships require conceptual support.

In the force unit, the crucial contrast is between an object's velocity and its acceleration: a zero resultant force implies zero acceleration, not necessarily zero velocity. EBDL is designed to surface and discuss this distinction through observable phenomena, force diagrams, and verbal justification. The anticipated pedagogical advantage is greater coherence among students' descriptions, diagrams, and equations; the present evaluation directly measures only written responses on the assessment.

2.2. *Experiential Learning as a Foundation*

Kolb's experiential cycle comprises four theoretical processes: concrete experience, reflective observation, abstract conceptualization, and active experimentation (Kolb, 2015). In this article, these are the theoretical inputs to EBDL, not an alternative four-phase description of the final model. The EBDL instructional syntax contains five phases with an additional explicit reflection-and-transfer closure.

Experiential tasks can make a force interaction salient, but completing a task does not demonstrate understanding of Newtonian relationships. EBDL therefore requires a traceable explanation at each transition: a prediction, a comparison of prediction and observation, a resultant-force argument, a contextual application, and a transfer statement (Morris, 2020; Savinainen & Scott, 2002).

2.3. Deep Learning as an Instructional Principle

Here, educational deep learning means explanation, evidence evaluation, integration across representations, and transfer to a new situation. Mindful learning is operationalized as deliberate attention and prediction; meaningful learning as reconciliation between observations and the scientific account; joyful learning as achievable, collaborative application without sacrificing conceptual rigor (Fullan et al., 2018). These terms describe intended learning conditions, not independently validated scales in this study.

For the topic of force, deep conceptual work includes assigning directions to forces, interpreting balanced and unbalanced interactions, explaining acceleration from net force, and distinguishing acceleration from velocity. Each form of reasoning aligns with the test blueprint in Appendix A and provides a specific learning target beyond general engagement or enjoyment (Hestenes et al., 1992; Docktor & Mestre, 2014).

2.4. The Five-Phase EBDL Model: Syntax and Distinctive Features

EBDL connects firsthand events to physics explanation by assigning each phase a distinct instructional function. Mindful Experiencing elicits predictions and directs observation; Reflective Meaning-Making identifies agreements and contradictions between initial ideas and evidence; Deep Conceptualization formalizes force diagrams, resultant force, and Newtonian justification; Joyful Knowledge Application practices reasoned solutions in new everyday tasks; and Transformative Reflection and Transfer asks learners to revisit their earlier reasoning and apply the principle in a modified context. The integration lies in this ordered chain of evidence and explanation, not merely in conducting experiments or using engaging activities.

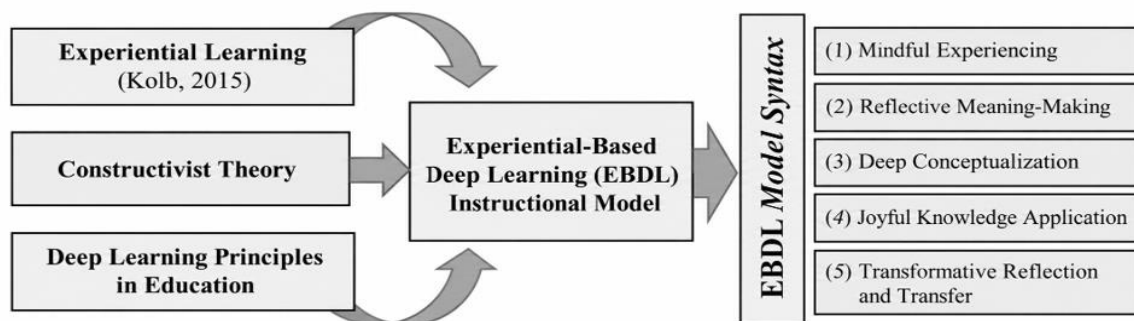


Figure 1. *Theoretical Foundations and Five Instructional Phases of EBDL*

The EBDL social system combines small-group observations and peer dialogue with explicit teacher-led scientific consolidation. The teacher reacts to an initial misconception by eliciting the student's reasoning, requesting evidence, inviting counter-explanations, and supplying targeted scaffolding when a scientifically coherent account has not emerged. The teacher does not treat every student statement as equally valid: explanations are checked against force diagrams, observed motion, and Newton's laws. Thus, structured inquiry is paired with explicit disciplinary guidance, whereas the Direct Instruction condition organizes the lesson primarily around teacher explanation followed by guided and independent exercise.

The support system comprises simple safe manipulatives (a toy cart or ball, a surface, and optional spring balances), short prediction and observation sheets, force-arrow diagrams, structured discussion prompts, and a transfer task. Its intended direct outcomes are accurate explanation of resultant forces, equilibrium, and force–acceleration relations as assessed by written tasks. Possible nurturant outcomes

include scientific argumentation, collaboration, metacognitive monitoring, and willingness to revise ideas; these latter outcomes were not measured. The difference from an ordinary demonstration is that every observation must be translated into a justified conceptual claim and then reconsidered in a different situation.

2.5. Force-Specific Cognitive Targets and Learning Mechanisms

The model targets a set of linked ideas in introductory mechanics: opposite forces must be combined with direction; a nonzero net external force is associated with acceleration; equal and opposite net forces acting on the same object imply no acceleration; and predictions about changing pushes must account for mass and relevant interactions. The assessment blueprint contains an integrated understanding/application task (Item 1, C2/C3), application of Newton's second law (Item 3, C3), analysis of changing motion and equilibrium (Items 2 and 4, C4), and evaluation of contextual pushing strategies (Item 5, C5). C2 is not measured by a standalone item, so a separate C2 effect cannot be inferred from the available data.

A typical EBDL reasoning chain begins when students predict whether unequal opposing pushes will move a table; students then observe or simulate the interaction, draw arrows to represent force magnitudes and directions, and explain the resulting acceleration using the sign of the net force. A subsequent task involving equal opposing pulls distinguishes equilibrium from motion, while a wardrobe-pushing scenario prompts evaluation under stated assumptions about direction, friction, and combined force. These examples specify the proposed mechanism of instruction; without item-level response coding or fidelity observations, they cannot be presented as documented mechanisms of the measured group difference.

EBDL's instructional logic differs from unstructured experiential activity because reflection is prompted explicitly, formal physics representations are required before application, and a final task tests whether students can reuse the principle beyond the original demonstration. It also differs from the specified Direct Instruction control condition in when students generate and test explanations: before rather than primarily after the teacher's exposition. Both conditions can include explanation, practice, and feedback; therefore, the comparison does not establish that only EBDL involves active cognition.

3. Method

3.1. Research Design and Study Stages

This quantitative quasi-experiment used a nonequivalent pretest–posttest control-group design. The investigation comprised three chronological stages: preparation (local needs analysis, force-topic alignment, EBDL instructional planning, instrument preparation, and expert appraisal); classroom implementation (pretesting, delivery of EBDL or Direct Instruction, and posttesting); and final analysis (score processing, between-group comparison, and interpretation). As the classes were not randomized at the student level, comparisons are associations under the study design and should not be interpreted as definitive causal effects (Creswell & Creswell, 2018). The primary reported contrast was between group posttest scores; no participant-level paired pretest–posttest dataset was supplied for an estimate of within-student change.

3.2. Setting, Sampling, and Participants

The study was conducted at SMP Negeri 1 Tengah in the second semester of the 2025/2026 academic year. The source manuscript reports purposive selection of 40 eighth-grade students: 20 in the EBDL group and 20 in the Direct Instruction control group. Selection considered teacher recommendations

and prior academic performance, but numeric baseline-equivalence diagnostics and individual pretest scores are not reported. The sampling frame, number of participating classrooms, age distribution, and allocation mechanism require confirmation from the original field records. Local selection and nonrandom assignment constrain generalizability and internal validity.

3.3. Intervention Design and Treatment Controls

The independent variable was the assigned instructional organization: the five-phase EBDL sequence versus the stated Direct Instruction sequence (presentation of objectives, teacher explanation, guided practice, independent practice, and evaluation). The measured dependent variable was post-instruction conceptual understanding of force on the essay assessment. For EBDL, mindful, meaningful, and joyful refer to how the learning tasks are structured rather than to separately measured outcomes. Both groups covered the same force-related content and completed the same testing instruments.

The source account reports that the same teacher, instructional goals, learning duration, core topic, and environment were used for both groups. These are intended controls rather than a guarantee of equivalent learning opportunities. A reproducible report should additionally document the actual number of meetings, minutes per meeting, materials used, teacher fidelity notes, and any deviations from the planned sequence. Those implementation-log details were not included in the supplied manuscript and cannot be reconstructed retrospectively.

Table 1. *Illustrative EBDL Lesson Cycle for the Force Topic*

EBDL phase (minutes)	Force-topic focus	Teacher reaction and support	Student cognitive activity	Observable product
Mindful Experiencing (10)	Opposing pushes and motion	Ask for initial prediction; demonstrate or guide a simple trial; withhold the final explanation	Predict directional motion; attend to force magnitudes; record observed movement and directional force representations.	Individual prediction and observation notes
Reflective Meaning- Making (15)	Prediction versus observed motion	Invite explanations, conflicting ideas, and evidence-based questions; make disagreement safe	Compare prediction with observation; verbalize why greater and opposing forces matter	Revised claim with supporting observation
Deep Conceptualization (25)	Vector sum, $F_{net} =$ ma , equilibrium	Scaffold arrow diagrams and Newtonian reasoning; challenge velocity– acceleration conflation	Represent force direction; derive net force; explain why balanced forces imply no acceleration	Force diagram and reasoned explanation
Joyful Knowledge Application (20)	Cart acceleration and pushing a wardrobe	Assign contextual cooperative tasks; provide formative feedback	Solve and justify transfer- near problems; examine assumptions about force and friction	Written answer with context and justification
Transformative Reflection and Transfer (10)	Changed mass or new opposing- force situation	Ask how reasoning changed; prompt application to a new situation	Reconsider initial model; articulate the transferable resultant-force principle	Reflection and exit transfer response

Preparation included literature-based model specification, activity planning and instrument development; implementation began with the same pretest for both groups, followed by the respective classroom treatments and posttest; the final stage involved descriptive and inferential analyses. The preparation stage did not include a separately reported empirical model-development trial, so this paper

evaluates an implemented instructional design rather than claiming full formative-development validation of a new model.

In EBDL lessons, the teacher first prompted predictions about opposing forces on a table or a ball's change in motion and guided direct observation (Mindful Experiencing). Students next compared initial claims with what they observed, explained differences, and questioned peers' accounts (Reflective Meaning-Making). The teacher then scaffolded force-arrow representations, the calculation of resultant force, and the distinction between equilibrium and acceleration (Deep Conceptualization). Students applied the reasoning to cart acceleration and collaborative pushing problems (Joyful Knowledge Application), before returning to an altered force situation and writing how their initial explanation had changed (Transformative Reflection and Transfer). The Direct Instruction group instead received explicit explanation, demonstrated examples, guided exercises, individual exercises, and evaluation. Table 1 gives a replicable lesson-cycle template; its times are illustrative and must not be mistaken for verified treatment durations.

3.4. *Assessment Instrument and Content Validity*

The outcome instrument contained five constructed-response questions aligned with the force-unit indicators and the revised Bloom taxonomy (Appendix A, Table 7). Item 1 combines basic interpretation and application of opposing forces (C2/C3); Item 3 measures application of $F_{\text{net}} = ma$ (C3); Items 2 and 4 address analysis of motion and equilibrium (C4); and Item 5 requests evaluative reasoning about pushing strategies (C5). Consequently, the blueprint contains one integrated C2/C3 item, one additional C3 item, two C4 items, and one C5 item, but no independent pure-C2 item. Five content dimensions were rated by two experts using a five-point scale: objective alignment, conceptual accuracy, cognitive-target alignment, representation consistency, and language. For each question, averaging the five criterion-specific Aiken V coefficients is equivalent to $V = \Sigma(r - 1)/[10(5 - 1)]$ over the ten ratings. Recalculation from the scores in Appendix B gives 0.875, 0.875, 0.875, 0.850, and 0.875, with overall $V = 0.870$ (Aiken, 1985); each exceeds the high-validity threshold of 0.80 used in the field account. The appended records provide numeric ratings but contain no verbatim validator comments or documented revision decisions, so such comments cannot be reported as observed findings. Although the manuscript refers to pilot testing and Cronbach's alpha, neither pilot-response scores nor an alpha estimate were supplied; instrument reliability therefore remains unverified from the available material (Taber, 2018).

3.5. *Data Processing and Statistical Analysis*

Both groups completed a pretest and posttest according to the research account; however, the available analysis tables report the posttest aggregate results only. Descriptive statistics summarize group central tendency and dispersion. Distributional checks used Shapiro–Wilk for each group (EBDL: $W = 0.893$, $p = 0.031$; control: $W = 0.889$, $p = 0.026$); these reported results motivated the nonparametric Mann–Whitney U comparison of independent posttest scores at $\alpha = 0.05$.

The effect-size statistic was r with $N = 40$. Since individual scores, pretest statistics, and item-level responses were not available for the reported analysis, baseline-adjusted effects, paired gains, confidence intervals, category-specific Mann–Whitney tests, and qualitative coding of students' common errors cannot be reliably reconstructed. No item-level significance statements are made.

The group distributions were reported to depart from normality on Shapiro–Wilk testing. Mann–Whitney U was accordingly used for the reported independent-group posttest contrast. A significant U test indicates a difference in rank distributions, not automatically a difference in mean improvement or a causal treatment effect. Any item-level comparisons should use recorded item scores, describe the scoring

rubric, and consider multiple-testing adjustment; analysis of C2 separately would require additional stand-alone indicators because Item 1 combines C2 and C3.

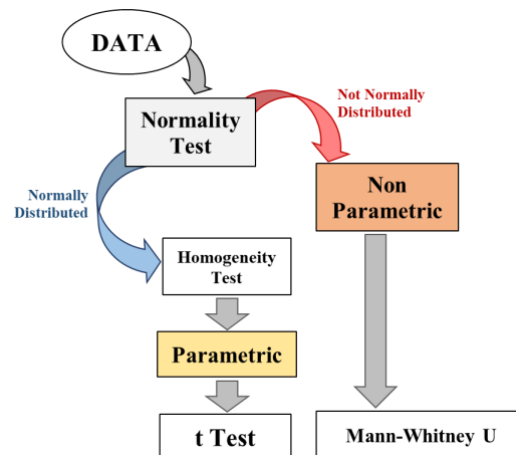


Figure 2. Statistical Decision Flow for Independent-Group Comparisons

Table 2. Criteria for Effect Size r ($Z/\sqrt{N}$)

r Value	Effect Category	Practical Interpretation in Educational Research
$0.00 \leq 0.10$	Very Small	The effect is practically insignificant or negligible.
$0.10 \leq 0.30$	Small	The intervention has an effect, but its impact is limited and often context-dependent.
$0.30 \leq 0.50$	Moderate	The educational effect becomes meaningful and worthy of consideration for instructional interventions.
≥ 0.50	Large	The intervention has a strong effect and is considered highly effective pedagogically.

The inferential null hypothesis stated that the posttest score distributions for the EBDL and control groups did not differ; the alternative hypothesis stated that they differed. The alpha level was 0.05, and a two-sided Mann-Whitney U test was used.

$$H_0 : \mu_E = \mu_C$$

$$H_1 : \mu_E \neq \mu_C$$

The groups are denoted E (EBDL) and C (Control); no Indonesian initial is used in the statistical hypotheses. Statistical significance was judged from the two-sided Mann-Whitney p value. This primary analysis addresses posttest differences rather than the magnitude of students' individual pretest-posttest changes.

4. Result

4.1. Descriptive Statistical Analysis

Table 3. Descriptive Statistical Analysis Data

Statistic	Experiment Group	Control Group
Mean	85.85	81.30
Median	87.00	83.50
Variance	30.35	49.06
Std. Deviation	5.51	7.00
Minimum	74.00	71.00
Maximum	93.00	91.00
Range	19.00	20.00

The post-instruction EBDL group had a mean score of 85.85, compared with 81.30 in the control group, a descriptive difference of 4.55 points (Table 3). The median was also higher in the EBDL group (87.00 versus 83.50). Together, these summaries indicate that the experimental group's distribution

tended toward higher total scores in this sample. Because no initial-score comparison is tabulated, this difference is not itself an estimate of learning gain.

Scores in the EBDL group were less dispersed ($SD = 5.51$; variance = 30.35) than those in the Direct Instruction group ($SD = 7.00$; variance = 49.06), suggesting more concentrated posttest performance. This pattern is compatible with the hypothesis that repeated prediction, discussion, force-diagram scaffolding, and transfer opportunities provide multiple entry points for understanding. However, the current dataset records neither fidelity to those activities nor the cognitive steps used in individual solutions, so the explanation remains pedagogically plausible rather than empirically established (Hake, 1998; Savinainen & Scott, 2002).

Both groups contained relatively high posttest scores, and their observed score ranges overlapped substantially (EBDL: 74.00–93.00; control: 71.00–91.00). Therefore, the difference is best interpreted as a moderate shift in the overall observed distribution, not as evidence that all EBDL students outperformed every control student or that a particular misconception disappeared. The more compact EBDL distribution does not, on its own, establish greater conceptual uniformity across individual force subtopics.

The five questions encompass distinct kinds of mechanics reasoning: vector combination, motion interpretation, applying Newton's law, equilibrium analysis, and evaluating a pushing strategy. A theory-based interpretation would associate Deep Conceptualization most directly with the representation and analysis tasks and Transformative Reflection and Transfer with evaluative contextual reasoning. These associations describe intended instruction-to-assessment alignment, not observed item-specific effects. Physics education research stresses that aggregated scores can conceal persistent conceptual difficulties (Hestenes et al., 1992; Docktor & Mestre, 2014).

Thus, the descriptive evidence supports higher posttest total conceptual-understanding scores for the EBDL group in this local comparison, while the available analysis is insufficient for concluding which cognitive domain improved most, whether the effect generalized to unfamiliar tasks, or which EBDL phase produced the score difference.

16

4.2. Inferential Statistical Analysis

4.2.1. Reported Assumption Checks

The Shapiro–Wilk results in Table 4 indicate departures from the normality assumption for the reported posttest scores in both groups ($p = 0.031$ and $p = 0.026$). These results explain the choice of the rank-based Mann–Whitney U comparison described in the Method.

21

Table 4. Results of the Shapiro–Wilk Normality Test

Group	Statistic	df	Sig.
Experiment	0.893	20	0.031
Control	0.889	20	0.026

Normality testing served as an analysis-selection check rather than an independent educational outcome. Its result cannot identify a particular pedagogical mechanism, nor can it substitute for a baseline equivalence assessment of the nonequivalent groups.

4.2.2. Hypothesis Testing

A two-sided Mann–Whitney U test compared the available post-instruction conceptual-understanding scores across the independent EBDL and control groups. The primary significance threshold was 0.05. The test examines rank-based between-group differences; it does not assess within-student pretest-to-posttest change.

24 **Table 5.** *Results of the Mann–Whitney U Test for Students' Science Conceptual understanding*

Statistic Test	Value
Mann–Whitney U	124.00
Z	−2.06
Asymp. Sig. (2-tailed)	0.039

The reported test returned $U = 124.00$, $Z = -2.06$, and two-sided $p = 0.039$ (Table 5). The negative Z sign reflects the rank convention; its magnitude is consistent with the reported p value. At $\alpha = 0.05$, the posttest distributions differed statistically.

Together with the descriptive results, this provides evidence of a between-group difference favoring EBDL on the total force-concept assessment. Owing to purposive sampling, the lack of baseline-adjusted analysis, and the absence of detailed fidelity records, the observed difference should not be treated as isolated proof that the EBDL model caused the outcome. The standardized rank-test effect size was calculated using,

$$r = \frac{|Z|}{\sqrt{N}}$$

with N representing the total number of tested students. Substituting the reported standardized test statistic gives.

$$r = \frac{|Z|}{\sqrt{N}} = \frac{2.06}{\sqrt{40}} = \frac{2.06}{6.325} = 0.326$$

The result is in the moderate range of the stated r interpretation criteria. The estimate summarizes the strength of the group contrast in the present sample; it is not a direct measure of the effectiveness of individual EBDL phases or of durable conceptual change.

Table 6. *Results of the Effect Size Calculation*

Parameter	Value
Z	−2.06
N	40
Effect Size (r)	0.326
Category	Medium

4.3. Cognitive-Process Coverage and Limits of Item-Level Inference

The assessment blueprint (Appendix A, Table 7) maps the opposing-force problem to an integrated C2/C3 process, the cart calculation to C3, the motion and balanced-force tasks to C4, and the wardrobe decision to C5. These task demands justify a domain-specific interpretation of the instructional design but are not sufficient for an empirical ranking of cognitive gains. The present analysis includes only total-score summaries and one omnibus Mann–Whitney test; no participant-by-item scoring matrix is available. Consequently, Mann–Whitney U tests for C3, C4, and C5, and examination of response errors on Item 4, remain analyses requiring the original scored responses. C2 should not be tested as a standalone construct using Item 1 alone without a validated subscore. This limitation is reported here instead of presenting imputed item-level statistics or invented quotations from student work.

5. Discussion

5.1. Interpretation of the Findings

The observed between-group advantage in total force-concept scores is consistent with the pedagogical rationale for combining prediction, direct observation, explicit representations, and reflection. Unlike a purely procedural demonstration, the EBDL sequence deliberately asks learners to connect a physical

event with a claim about net force and then justify that claim in a new context. Evidence from interactive engagement in mechanics suggests that opportunities to articulate, test, and revise explanations can support conceptual learning (Hake, 1998; Savinainen & Scott, 2002). The present result aligns with that broader finding but does not identify a uniquely causal EBDL component.

A central feature of the model is the transition from intuitive movement descriptions to Newtonian explanations. The activity involving unequal opposing pushes requires students to attend to force direction; the equal-force situation makes the distinction between equilibrium and acceleration discussable. These contexts are especially relevant because force-concept research documents persistent confusion between the presence of force, velocity, and acceleration (Hestenes et al., 1992; Chrysostomou et al., 2024). The instructional design is therefore aligned with an established conceptual difficulty rather than with deep learning understood only as an abstract pedagogical label.

Reflective Meaning-Making and Deep Conceptualization are plausible mechanisms for this alignment: verbal predictions create a reference point, while diagrams and resultant-force arguments offer a way to evaluate whether a later explanation is scientifically consistent. Research in physics education emphasizes the importance of linking representations and testing intuitive explanations, and studies of interactive learning similarly associate such opportunities with conceptual gains (Docktor & Mestre, 2014; Banda & Nzabahimana, 2021). Nevertheless, the available aggregate scores cannot verify which phase contributed more strongly or determine whether particular initial misconceptions were replaced.

The study instrument offers a meaningful route for deeper analysis without supporting an invented interpretation of student responses. The equilibrium item asks learners to identify the consequences of balanced forces, while the contextual evaluation item asks whether claims about combined pushes are justified. Original item responses could reveal whether students confuse no acceleration with no motion or overlook assumptions about friction. Until those responses are coded and compared between groups, such errors must be discussed as documented possibilities in the physics-education literature, not as observed characteristics of participants in this study (Hestenes et al., 1992; Savinainen & Scott, 2002).

The EBDL sequence can be understood as a force-specific operationalization of experiential learning rather than a wholesale replacement of established models. Its strength is the planned succession of eliciting a claim, confronting it with an experience, formalizing the physics explanation, and testing transfer. Kolb's theory motivates learning from experience, while research on conceptual understanding in physics motivates the explicit representations and targeted questions (Kolb, 2015; Docktor & Mestre, 2014). This integration is a theoretical and design contribution; superiority over all other active-learning approaches has not been tested.

At the same time, the comparison should not be read as evidence that direct explanation or teacher guidance is intrinsically counterproductive. Both treatment conditions included structured instruction and assessment, and EBDL itself depends on carefully timed explanatory scaffolding. The distinctive feature lies in the opportunities for students to generate and revise accounts before and after formalization. This interpretation avoids equating a statistically favorable posttest result with proof of every proposed mechanism.

A practical strength of the study is that the EBDL design can be specified using concrete junior secondary force tasks, teacher response rules, low-cost teaching supports, and explicit cognitive targets. This makes the instructional proposal open to replication and falsification. Its empirical strength is a controlled-topic classroom comparison using a common assessment; its evidential boundary is the

absence of item-level, process, and baseline-adjusted analyses. These distinctions are essential to a proportionate interpretation of model effectiveness.

5.2. Implications for Physics Education

For physics teachers, the model offers a practical planning logic for connecting ordinary push-and-pull experiences with disciplinary reasoning about net force and acceleration. It encourages teachers to plan prediction prompts, observable checks, explicit force diagrams, targeted feedback, and reflection rather than treating hands-on participation as an endpoint. This could be adapted to electricity, energy, and other physics topics in which everyday intuition may conflict with a scientific representation; such adaptation should be tested rather than assumed to work automatically.

The broader international relevance is conceptual and methodological: physics curricula in diverse settings face similar difficulties with intuitive mechanics reasoning, and a transparent phase-by-phase protocol permits cross-contextual evaluation against established active-learning approaches. Replication would be strengthened by publishing lesson plans, scoring rubrics, implementation-fidelity indicators, and deidentified item-level datasets. These resources could help determine whether EBDL's deliberate integration adds value beyond its constituent practices and whether its transfer tasks support durable scientific understanding.

5.3. Research Limitations and Future Directions

Interpretation is limited by purposive sampling at a single school, the absence of documented allocation details and baseline-adjusted comparisons, and the narrow focus on introductory force. Local classroom norms and culturally familiar contexts may affect participation and the interpretation of everyday examples, restricting transfer to other educational settings. The analytic dataset contained no original student response sheets or scored item-level data, so the study cannot demonstrate remediation of a specific misconception or an advantage in a particular cognitive domain.

The teaching description also lacks original session logs, an implementation-fidelity measure, verbatim validator feedback, and a documented reliability estimate for the test. These omissions prevent confident attribution of the group difference to particular EBDL phases and constrain replication of the historical intervention. Future investigations should preregister cognitive-process hypotheses, document teacher enactment, report participant-level pretest and posttest outcomes, code common force-reasoning errors, and compare EBDL with active-learning controls across several schools.

A further priority is to test retention and transfer using conceptually parallel tasks and to evaluate whether learners improve specifically in reasoning about force interactions, equilibrium, and the relationship between net force and acceleration. Such evidence would distinguish a general posttest advantage from a demonstrated conceptual mechanism and clarify the conditions under which the EBDL sequence can contribute to physics education.

6. Conclusion

This classroom study found higher total post-instruction conceptual-understanding scores on introductory force tasks among students taught with the EBDL sequence than among students in the locally selected Direct Instruction group. The reported rank-based comparison supports a between-group difference, but the available summaries cannot establish within-student learning gains, removal of particular misconceptions, or the independent effect of any EBDL phase.

The principal contribution is a more explicit force-specific EBDL instructional architecture: students move from prediction and observation to evidence-based reflection, representation of net force, contextual problem solving, and conceptual transfer. The design identifies concrete teacher responses, supports, and intended cognitive outcomes, making the approach more amenable to systematic replication and evaluation than a purely generic account of experiential or deep learning.

Future multi-site studies should include baseline-adjusted analyses, participant-by-item responses, validation of scoring reliability, fidelity observations, and delayed transfer assessments. These steps are necessary to determine which aspects of force reasoning are most responsive to EBDL and whether the promising local posttest pattern generalizes.

Authors Contribution

Christo Vincentius Goni: Conceptualization, Methodology, Writing – original draft, Formal analysis.

Orbanus Naharia: Conceptualization, Methodology, Writing – review & editing, Supervision.

Marianus Marianus: Methodology, Data curation, Formal analysis, Writing – review & editing, Supervision. **Kenny Setiawan Lahope:** Methodology, Formal analysis, Writing – review and editing, Supervision. **Darius Ladbam Lihu:** Data curation, validation, writing – review & editing.

Ethical statement

This study was conducted in regular classroom teaching settings involving eighth-grade students at SMP Negeri 1 Tengah and was classified as minimal-risk educational research. Permission to conduct the study and to collect data in the participating classes was obtained from the relevant institutional authorities through an official research permission letter. Participation was voluntary, and informed consent was obtained from all participants. No sensitive personal data was collected. All responses were anonymized and used for research purposes.

Declaration of AI Use

The authors used ChatGPT (OpenAI) to improve text editing and grammar in this manuscript. Gemini AI was utilized to translate the text from Indonesian into English. All AI-assisted outputs were thoroughly reviewed and edited by the authors, who remain fully responsible for the accuracy, originality, and integrity of the final manuscript.

Conflict of Interest

The authors declare that there are no financial or non-financial conflicts of interest that could have influenced the work reported in this manuscript. All authors have reviewed and approved this statement.

Supplementary Materials and Data Availability

No public repository is currently available for the dataset. However, the instruments and key data summaries used in this study can be obtained from the corresponding author upon reasonable request. Any data shared will be anonymized and provided in accordance with the ethical approval and institutional regulations.

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>

- Andayanie, L. M., Adhantoro, M. S., Purnomo, E., & Kurniaji, G. T. (2025). Implementation of deep learning in education: Towards mindful, meaningful, and joyful learning experiences. *Journal of Deep Learning*, 1(1), 1–12. <https://doi.org/10.23917/jdl.v1i1.11157>
- Banda, H. J., & Nzabahimana, J. (2021). Effect of integrating physics education technology simulations on students' conceptual understanding in physics: A review of literature. *Physical Review Physics Education Research*, 17(2), Article 023108. <https://doi.org/10.1103/PhysRevPhysEducRes.17.023108>
- Biggs, J., & Tang, C. (2011). Teaching for quality learning at university (4th ed.). Open University Press.
- Chrysostomou, A., Cornell, A. S., & Naylor, W. (2024). A three-year comparative study of dominant misconceptions among first-year physics students at a South African university. *Physics Education*, 59(1), Article 015036. <https://doi.org/10.1088/1361-6552/ad14ea>
- Creswell, J. W., & Creswell, J. D. (2018). Research design: Qualitative, quantitative, and mixed methods approaches (5th ed.). Sage.
- Docktor, J. L., & Mestre, J. P. (2014). Synthesis of discipline-based education research in physics. *Physical Review Special Topics—Physics Education Research*, 10(2), Article 020119. <https://doi.org/10.1103/PhysRevSTPER.10.020119>
- Feri, M. (2025). Penerapan pembelajaran mendalam melalui joyful, meaningful, dan mindful learning: Tinjauan literatur. *IBTIDA: Jurnal Kajian Pendidikan Dasar*, 5(2), 145–163. <https://doi.org/10.33507/ibtida.v5i2.3495>
- Feriyanto, F., & Anjariyah, D. (2024). Deep learning approach through meaningful, mindful, and joyful learning: A library research. *Electronic Journal of Education, Social Economics and Technology*, 5(2), 208–212. <https://doi.org/10.33122/ejeset.v5i2.321>
- 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.
- 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>
- Hattie, J. (2023). Visible learning: The sequel. Routledge.
- Hestenes, D., Wells, M., & Swackhamer, G. (1992). Force concept inventory. *The Physics Teacher*, 30(3), 141–158. <https://doi.org/10.1119/1.2343497>
- Koç, A., & Kanadlı, S. (2025). Effect of interactive learning environments on learning outcomes in science education: A network meta-analysis. *Journal of Science Education and Technology*, 34(4), 681–703. <https://doi.org/10.1007/s10956-025-10202-7>
- Kolb, D. A. (2015). Experiential learning: Experience as the source of learning and development (2nd ed.). Pearson.
- Kulgemeyer, C., & Wittwer, J. (2023). Misconceptions in physics explainer videos and the illusion of understanding: An experimental study. *International Journal of Science and Mathematics Education*, 21, 417–437. <https://doi.org/10.1007/s10763-022-10265-7>
- Marjana, R., Samsudin, A., Saepuzaman, D., & Rusdiana, D. (2026). Misconception in physics learning: A systematic review of the research trend, misconception profile, diagnosis, and remediation

- strategies. *Jurnal Pendidikan Fisika*, 14(1), 240–262. <https://doi.org/10.26618/7wdrxx78>
- 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>
- Morris, T. H. (2020). Experiential learning: A systematic review and revision of Kolb's model. *Interactive Learning Environments*, 28(8), 1064–1077. <https://doi.org/10.1080/10494820.2019.1570279>
- Nafi'ah, J., & Faruq, D. J. (2025). Conceptualizing deep learning approach in primary education: Integrating mindful, meaningful, and joyful. *Journal of Educational Research and Practice*, 3(2), 225–237. <https://doi.org/10.70376/jerp.v3i2.384>
- OECD. (2023). PISA 2022 assessment and analytical framework. OECD Publishing. <https://doi.org/10.1787/dfe0bf9c-en>
- OECD. (2023). PISA 2022 results (Volume I): The state of learning and equity in education. OECD Publishing. <https://doi.org/10.1787/53f23881-en>
- Piaget, J. (1977). The development of thought: Equilibration of cognitive structures. Viking Press.
- Rahmawati, Y., Mu'ti, A., Suyanto, S., & Herianingtyas, N. L. R. (2025). Pembelajaran mendalam: Transformasi pembelajaran menuju pendidikan bermutu. *Jurnal Penelitian Kebijakan Pendidikan*, 18(1). <https://doi.org/10.24832/jpkp.v18i1.1281>
- Ramadhan, A. (2025). Pengaruh meaningful, joyful, dan mindful learning sebagai pilar deep learning terhadap hasil belajar: Literature review. *Jurnal Pendidikan Tematik*, 6(2), 215–225. <https://doi.org/10.62159/jpt.v6i2.1736>
- Ratri, S. Y., Wagiran, W., Riyadi, A., Firdaus, F. M., & Delosa, J. (2026). Deep learning (Pembelajaran mendalam) dalam pendidikan Indonesia: Tinjauan sistematis teori dan praktik. *Jurnal Ilmu Pendidikan dan Pembelajaran*, 4(2), 176–191. <https://doi.org/10.58706/jipp.v4n2.p176-191>
- Retnawati, H. (2017). Validitas dan reliabilitas karakteristik butir (2nd ed.). Parama Publishing.
- Savinainen, A., & Scott, P. (2002). Using the Force Concept Inventory to monitor student learning and to plan teaching. *Physics Education*, 37(1), 53–58. <https://doi.org/10.1088/0031-9120/37/1/307>
- Siregar, T., Fauzan, A., Yerizon, Y., & Syafriandi, S. (2025). Designing mathematics teaching through deep learning pedagogy: Toward meaningful, mindful, and joyful learning. *Journal of Deep Learning*, 1(2), 85–97. <https://doi.org/10.23917/jdl.v1i2.11969>
- Syafi'i, A., & Darnanengsih. (2025). Pendekatan pembelajaran berbasis deep learning: Mindful learning, meaningful learning, dan joyful learning. *Al-Mumtaz: Jurnal Manajemen Pendidikan Islam*, 2(1), 1–15. <https://doi.org/10.47945/al-mumtaz.v2i1.199>
- Taber, K. S. (2018). The use of Cronbach's alpha when developing and reporting research instruments in science education. *Research in Science Education*, 48(6), 1273–1296. <https://doi.org/10.1007/s11165-016-9602-2>
- Wandi, W., Mardianti, F., Suwarma, I. R., & Liliawati, W. (2023). Theory and practice of conceptual understanding in physics education: A literature review and bibliometric analysis of recent decades. *Kasuari: Physics Education Journal (KPEJ)*, 6(2), 107–117. <https://doi.org/10.37891/kpej.v6i2.483>
- Widagdo, T. B. (2025). Pandangan konseptual pembelajaran mendalam menuju transformasi pendidikan. *Jurnal Cerdik: Jurnal Pendidikan dan Pengajaran*, 4(2). <https://doi.org/10.21776/ub.jcerdik.2025.004.02.05>
- Yardley, S., Teunissen, P. W., & Dornan, T. (2012). Experiential learning: Transforming theory into practice. *Medical Teacher*, 34(2), 161–164. <https://doi.org/10.3109/0142159X.2012.643264>

APPENDIX A. Research Instrument for Conceptual Understanding of Force

Table 7. Test Blueprint of the Research Instrument

No.	Indicator	C Level	Test Item	Answer Key
1	Students apply the concept of the resultant force of forces acting in the same direction and in opposite directions on an object.	integrated C2 and C3	 A table is pushed by Andi with a force of 80 N to the right. At the same time, Budi pushes the table with a force of 50 N to the left. Determine the magnitude and direction of the resultant force acting on the table.	Given: $F_1 = 80 \text{ N (right)},$ $F_2 = 50 \text{ N (left)}.$ Required: Determine the magnitude and direction of the resultant force acting on the table Solution: The resultant force is calculated as: $\Sigma F = 80 - 50 = 30 \text{ N}.$ The direction of the resultant force follows the direction of the greater force, which is to the right. Therefore, the resultant force acting on the table is 30 N to the right.
2	Students analyze the effect of force on changes in the motion of an object based on the provided figure.	C4	Observe the following figure.  Explain the changes that occur to the ball as a result of the applied force and analyze the relationship between the force and the motion of the ball.	The ball is initially at rest. The kick applies a force that causes the ball to move. Thus, the force changes the state of motion of the object from rest to motion. The greater the applied kicking force, the greater the change in the ball's velocity.
3	Students apply the principles of Newton's Second Law to everyday situations.	C3	A cart with a mass of 20 kg is pushed with a force of 60 N on a smooth horizontal surface. Calculate the acceleration of the cart.	Given: $m = 20 \text{ kg},$ $F = 60 \text{ N}$ Using Newton's Second Law $a = \frac{F}{m} = \frac{60}{20} = 3 \text{ m/s}^2$
4	Students analyze the equilibrium condition of forces acting on an object based on a force diagram.	C4	Observe the illustration of two men pulling a box shown below.  In the figure, a force of 40 N is applied from the left and a force of 40 N is applied from the right. Analyze the condition of the object. Does the object experience acceleration? Explain your answer based on the concept of resultant force.	Given: Force acting to the left = 40 N and Force acting to the right = 40 N Required: Determine the condition of the object. Does the object experience acceleration? Solution: The resultant force is calculated as: $\Sigma F = 40 - 40 = 0 \text{ N}.$ Since the resultant force is zero, the object is in a state of equilibrium. Therefore, the object does not experience acceleration. If the object is initially at rest, it will remain at rest. If it is already in motion, it will continue moving with constant velocity.
5	Students evaluate the application of	C5	Consider the following illustration.	The more effective method is when two students push together. The total

force in everyday life and provide scientific reasoning.



In your opinion, which method is more effective for moving a wardrobe? Evaluate it based on the concept of force and provide scientific justification.

force acting on the object becomes larger because the forces exerted by both students are combined. A greater resultant force produces greater acceleration, making the wardrobe easier to move. The evaluation may also consider the efficiency of effort and time.

APPENDIX B. Validation of Research Instruments

The validation process concerned the force-concept essay assessment (including resultant force, motion, equilibrium, and Newtonian application). Two raters provided five-point content ratings on the following criteria:

1. alignment of test items with learning objectives,
2. conceptual core or essential meaning of the concept,
3. suitability of items for measuring conceptual understanding indicators,
4. consistency in the use of representations (verbal, visual, and mathematical) within the items, and
5. language and writing quality.

The scale ranged from 1 (not valid) to 5 (highly valid). For each criterion, Aiken's V was calculated from both validators as $V = \frac{\Sigma(r-1)}{[2(5-1)]}$. An item-level summary was then obtained by averaging the five criterion-specific V values. Table 8 reports the corrected item coefficients from the original ratings, with a mean of 0.870. The retained validation record contains no transcribed validator comments or documented revision log, so qualitative validator feedback cannot be evaluated. No empirical internal-consistency coefficient was available for verification.

Table 8. *Validation of the Research Instrument*

Question Number	VALIDATOR 1					VALIDATOR 2					Aiken's V	Category
	Assessment Criteria					Assessment Criteria						
	(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)		
1	5	5	4	4	5	5	5	4	4	4	0.875	High validity
2	4	5	4	4	5	4	5	5	4	5	0.875	High validity
3	5	4	4	5	4	5	4	5	5	4	0.875	High validity
4	4	5	4	5	4	4	4	5	5	4	0.850	High validity
5	4	4	5	5	4	5	5	4	5	4	0.875	High validity

Across the five assessment items, the mean Aiken's V was 0.870, computed from the original two-validator, five-criterion ratings. The corrected validity coefficients represent expert agreement on content relevance, not evidence of inter-rater scoring reliability or proven sensitivity to treatment effects.