



The Application of the Discovery Learning Method in Teaching Zakat to Junior High School Students at Safinatussalam Islamic Boarding School

Muhammad Riza^{1*} Ilyas Fajar^{2*}

¹⁾ IAIN Takengon, rizajundana@gmail.com

²⁾ MTs Pondok Pesantren Safinatussalam, ilyasfajar01@gmail.com

DOI:



Copyright © 2023

Diajukan: 10/01/2026

Diterima: 13/02/2026

Diterbitkan: 14/03/2026

ABSTRAK

Zakat occupies a constitutive position in Islamic practice, yet its treatment in secondary religious instruction is frequently reduced to the memorisation of thresholds and ratios, producing computational fluency without moral or civic engagement. This study examines how the discovery learning method is enacted in zakat instruction and what forms of learning it generates in a setting where the subject matter is doctrinally fixed rather than empirically negotiable. Using a qualitative single-case design with embedded units, the study was conducted over one academic semester at Safinatussalam Islamic Boarding School in Indonesia, involving one fiqh teacher, twenty-four Grade VIII students, six focal learners, and two institutional informants. Data were generated through eight non-participant lesson observations, semi-structured and stimulated-recall interviews, one focus group discussion, and the analysis of 96 student work artefacts, and were interpreted through reflexive thematic analysis supported by NVivo 14. Findings show that the teacher relocated the object of discovery from the ruling itself to the reasoning that connects a ruling to a contemporary case, producing a pattern the study terms bounded discovery; that guidance was calibrated rather than withdrawn; and that gains in case-based reasoning and moral disposition outpaced gains in computation. Institutional time pressure, examination logic, and heterogeneous prior knowledge constrained the deepest phases of inquiry. The study contributes an epistemically differentiated account of discovery learning that specifies what is discoverable in a revealed-knowledge domain, extends inquiry research beyond its science and mathematics heartland into pesantren education, and offers teachers a defensible design logic for teaching normative content without abandoning learner agency.

Keywords: discovery learning; zakat education; Islamic boarding school; fiqh pedagogy; guided inquiry; qualitative case study; religious literacy

ABSTRACT

Zakat menempati posisi konstitutif dalam praktik keislaman, namun pengajarannya pada jenjang menengah kerap tereduksi menjadi hafalan nisab dan kadar, sehingga menghasilkan kecakapan hitung tanpa keterlibatan moral dan sosial. Penelitian ini mengkaji bagaimana metode discovery learning diterapkan dalam pembelajaran zakat serta bentuk pembelajaran apa yang dihasilkannya pada bidang kajian yang materinya bersifat tetap secara doktrinal dan tidak dapat dinegosiasikan secara empiris. Dengan desain studi kasus tunggal terpancang berpendekatan kualitatif, penelitian dilaksanakan selama satu semester di Pondok Pesantren Safinatussalam, Indonesia, melibatkan satu guru fikih, dua puluh empat siswa kelas VIII, enam siswa fokus, dan dua informan kelembagaan. Data dikumpulkan melalui delapan observasi non-partisipan, wawancara semi-terstruktur dan stimulated recall, satu diskusi kelompok terarah, serta analisis 96 artefak kerja siswa, kemudian dianalisis dengan analisis tematik reflektif berbantuan NVivo 14. Hasil penelitian menunjukkan bahwa guru memindahkan objek penemuan dari hukum itu sendiri kepada penalaran yang menghubungkan hukum dengan kasus kontemporer, sebuah pola yang disebut bounded discovery; bimbingan dikalibrasi dan bukan ditiadakan; serta peningkatan penalaran kasus dan disposisi moral melampaui peningkatan kemampuan berhitung. Keterbatasan waktu institusional, logika ujian, dan keragaman pengetahuan awal membatasi tahap penemuan yang paling mendalam. Penelitian ini menyumbangkan penjelasan discovery learning yang terdiferensiasi secara epistemik, memperluas kajian inkuiri ke ranah pendidikan pesantren, dan menawarkan logika desain pembelajaran normatif yang tetap menjaga agensi peserta didik.

*Korespondensi Author: Ilyas Fajar, ilyasfajar01@gmail.com

I. INTRODUCTION

Zakat is not merely one obligation among others in Islamic practice; it is the institutional mechanism through which a confessional community converts private wealth into public welfare, and its effectiveness therefore depends on a population that understands it rather than merely performs it. Contemporary evidence from the Muslim world consistently locates knowledge, rather than piety alone, at the centre of zakat behaviour. Studies of compliance intention across entrepreneurial and salaried populations show that knowledge of the obligation moderates the relationship between religious motivation and actual payment, so that religiosity unaccompanied by understanding does not reliably translate into fulfilment (Sadallah et al., 2023). Research on educational background and compliance reaches a convergent conclusion: formal exposure to the substance of the obligation, and not merely to its existence, predicts whether individuals discharge it correctly (Abdullah & Sapiei, 2018). If the determinants of adult behaviour are cognitive as much as devotional, then the school years in which those cognitive foundations are laid deserve far more analytical attention than they have received. The classroom, in other words, is a site of zakat policy, even though it is rarely described as one.

The recent literature on zakat education, however, has developed asymmetrically. Investigations of literacy and awareness among student populations have begun to appear, and they report that literacy exerts a measurable influence on students' willingness to discharge the obligation while awareness alone proves a weaker predictor (Malik et al., 2024). What such studies establish is that a knowledge deficit exists and that it matters; what they do not establish is how instruction might remedy it, because their designs are survey-based and their unit of analysis is the individual respondent rather than the lesson. The pedagogical literature in Islamic education displays the mirror-image limitation. A substantial body of work now documents the persistence of transmissive teaching in religious subjects, the difficulty teachers experience in translating value commitments into classroom activity, and the institutional pressures that reproduce lecture-and-recitation formats even where policy encourages otherwise (Alabdulhadi & Alkandari, 2024; Kosim et al., 2023). Analyses of Islamic schooling internationally identify a comparable pattern, arguing that pedagogical transformation in Muslim educational institutions has been slower than curricular reform because teacher formation, assessment regimes and inherited conceptions of authority remain largely untouched (Memon et al., 2023). Between these two literatures lies an unoccupied space: research that treats zakat not as an object of attitudinal measurement and not as a generic instance of religious content, but as a specific instructional problem with its own epistemic character.

Evidence on the instructional alternative is, by contrast, unusually strong, though drawn from other subject areas. Synthetic reviews of inquiry-oriented instruction report consistent positive effects on conceptual understanding and higher-order thinking, with meta-analytic estimates in science education clustering in the small-to-moderate range and rising when the outcome measured is reasoning rather than recall (Antonio & Prudente, 2024; Alarcón et al., 2023). Comparable syntheses in Indonesian secondary and vocational settings report the same directional pattern for higher-order thinking outcomes, indicating that the effect is not confined to Western instructional cultures (Agussuryani et al., 2022). The determining variable in this body of work is not the presence of discovery but the calibration of guidance. The canonical meta-analysis on the question demonstrates that inquiry produces reliable gains when learners receive structured support, and that the effect diminishes or reverses when support is withdrawn (Lazonder & Harmsen, 2016). Subsequent systematic review of fully open inquiry confirms the asymmetry, finding that the least structured variants yield the most conditional outcomes and depend heavily on learner expertise and task familiarity (Muhamad Dah et al., 2024). Parallel evidence from Asian classrooms, where

instructional cultures share features with Indonesian settings, indicates that active approaches outperform transmissive ones while remaining sensitive to class size, examination pressure and teacher preparedness (Ting et al., 2023). Complementary work shows that the sequencing of exploration and instruction is itself consequential: problem-solving that precedes explicit teaching produces deeper conceptual outcomes than the reverse ordering, provided the subsequent consolidation is deliberate (Sinha & Kapur, 2021). Studies of structured and confirmation inquiry add a motivational dimension, reporting improvements in achievement motivation even where achievement gains are modest (Toma, 2022), and technology-mediated variants extend the pattern into digitally supported secondary classrooms (Liu et al., 2021; Marion et al., 2023). Taken together, this evidence supports a guided, not a *laissez-faire*, reading of discovery.

The theoretical apparatus that best accommodates these findings combines constructivist accounts of concept formation with cognitive-architectural constraints on how much unassisted processing novices can sustain. Discovery learning in its defensible form treats the learner as an active reconstructor of meaning who advances through stimulation, problem formulation, data gathering, data processing, verification and generalisation. Cognitive load theory supplies the necessary corrective, specifying that biologically secondary knowledge must be acquired under conditions that protect limited working memory, and that guidance functions not as a concession to weak learners but as a structural requirement of novice cognition (Sweller, 2022). The productive tension between these positions has been resolved empirically rather than doctrinally: exploration is valuable precisely because it prepares learners to profit from subsequent instruction, and its value collapses when consolidation is omitted (Sinha & Kapur, 2021). This synthesis is appropriate to zakat instruction for a reason that is specific to the domain. Zakat comprises both fixed elements, such as the categories of eligible recipients and the thresholds and proportions established in the sources, and variable elements, such as whether a particular contemporary income stream falls within a given category. The former cannot be discovered in any meaningful sense; the latter can be reasoned toward. A theory of guided discovery that distinguishes between what is given and what is inferable is thus not an optional refinement in this domain but a precondition of coherent instructional design.

Four gaps follow from this reading of the field, and each is consequential. The first is methodological. Research on discovery learning in Islamic religious education has been dominated by short-cycle classroom action research and pre-test–post-test comparisons in which the reported outcome is an aggregate score and the instructional process itself remains a black box; the resulting evidence establishes that something worked without specifying what was done, which severely limits replication and transfer. Process-level qualitative evidence, capable of describing how teachers adapt inquiry phases under real constraints, is correspondingly scarce. The second gap is contextual. The empirical base for inquiry-oriented instruction is concentrated in science and mathematics classrooms in conventional schools (Alarcón et al., 2023; Antonio & Prudente, 2024), whereas the *pesantren*, an institution with distinctive authority relations, textual traditions and residential organisation, remains largely outside this evidence base even though it educates a substantial share of Indonesian adolescents (Kawakip, 2023). Existing *pesantren* scholarship has concentrated on institutional identity, leadership and value transmission (Jusubaidi et al., 2024; Karim et al., 2023) rather than on the micro-mechanics of subject teaching, and the few studies that do examine active methods in these institutions report implementation at the level of the whole approach without disaggregating which components of the content were opened to students and which were not (Rohman, 2022). The third gap is empirical. Zakat research remains overwhelmingly oriented toward adult payers and institutional performance (Abdullah & Sapiei, 2018; Sadallah et al., 2023), so that the formative period during which zakat literacy is actually built has generated only a thin body of classroom evidence, and adolescents are studied as prospective payers rather than as present learners (Malik et al., 2024). The fourth gap is theoretical. Constructivist discovery models presuppose a knowledge domain in which learners can, in principle, reconstruct the target concept from evidence. Revealed normative content violates that presupposition. No available framework specifies which components of a doctrinally fixed subject are legitimately discoverable, and in the absence of such a specification teachers are left either to abandon inquiry as theologically inappropriate or to apply it indiscriminately in ways that misrepresent the epistemic status of the material. These gaps matter jointly rather than severally: without process evidence the

methodological question cannot be answered, without pesantren data the contextual claim cannot be tested, and without an epistemic account of discoverability neither can be interpreted.

The present study addresses this configuration through three contributions. Conceptually, it advances and evidences the notion of bounded discovery, an account in which the object of inquiry is displaced from the ruling to the reasoning that links an established ruling to an unfamiliar case, thereby preserving the authority of the source while restoring genuine cognitive work to the learner. This reframing supplies the domain sensitivity that generic constructivist models lack and offers a criterion, rather than an intuition, for deciding what may be left for students to find. Methodologically, the study combines lesson observation with stimulated-recall interviewing and systematic analysis of student work artefacts, so that claims about learning are anchored in what students produced rather than in what participants recalled, an integration seldom applied to religious-subject classrooms. Contextually, it locates the inquiry inside a boarding institution in which the same students encounter both classical text study and national curriculum instruction, permitting analysis of how an imported pedagogical method is refracted through an indigenous epistemic tradition rather than assuming it arrives intact.

This study aims to describe how the discovery learning method is enacted in the teaching of zakat to junior high school students at Safinatussalam Islamic Boarding School; to characterise the forms of learning, in conceptual, procedural and dispositional terms, that this enactment generates; and to identify the institutional, epistemic and pedagogical conditions that enable or constrain its implementation. Three questions organise the inquiry. First, how do teacher and students enact the phases of discovery learning when the subject matter is doctrinally fixed? Second, what evidence of conceptual understanding, applicative reasoning and moral disposition emerges from student work and talk across the instructional sequence? Third, which contextual conditions mediate the fidelity and depth of implementation?

II. METHODS

This study adopted an interpretive qualitative orientation and a single-case design with embedded units of analysis, a configuration selected because the object of inquiry is a bounded contemporary phenomenon, the enactment of a specific method in a specific institution, that cannot be separated from its setting without losing its meaning. The case was defined as the teaching of zakat within the fiqh curriculum of Grade VIII during one academic semester at Safinatussalam Islamic Boarding School, an Indonesian residential institution combining classical text instruction with the national junior secondary curriculum. Three embedded units were specified in advance: the instructional sequence as designed and enacted, the learning trajectories of six focal students, and the institutional conditions surrounding the classroom. The rationale for a single case rather than a comparative design was analytical rather than pragmatic. The study seeks conceptual generalisation to theory concerning discoverability in normative domains, not statistical generalisation to a population of schools, and depth of contextual description is therefore the primary source of inferential warrant. Consistent with this logic, the researcher occupied a non-participant observer role throughout and had no responsibility for lesson planning, instruction or assessment.

Participants were selected purposively using criterion sampling supplemented by maximum-variation sampling at the focal-student level. The teacher participant was the fiqh instructor responsible for the zakat unit, selected on three criteria: a minimum of five years of teaching experience in the institution, prior formal exposure to discovery learning through in-service training, and willingness to permit observation of unmodified practice. The whole-class unit comprised twenty-four Grade VIII students, all of whom consented to observation and artefact collection. Six focal students were then selected to maximise variation in prior fiqh attainment, in duration of residence in the pesantren, and in classroom participation as judged by baseline observation, on the reasoning that discovery-based instruction is theorised to affect learners differently according to prior knowledge (Sweller, 2022) and that a homogeneous focal sample would suppress precisely the variation of interest. Two institutional informants, the deputy head for curriculum and a senior residential instructor, were included to address the third research question. Recruitment continued until the analytic categories generated by successive observations and interviews ceased to yield new dimensions, an information-power judgement rather than a fixed numerical target (Braun & Clarke, 2019).

Four instruments were used. An observation protocol structured around the six phases of discovery learning recorded, for each phase, the teacher moves, the student activity, the form and timing of guidance, and the duration allocated, together with an open-ended running record to capture unanticipated events. A semi-structured interview guide for the teacher and institutional informants addressed instructional rationale, perceived constraints and adaptation. A stimulated-recall protocol for focal students used short video segments and the students' own written work as prompts, a design choice intended to reduce retrospective rationalisation, which is a recognised threat in interview-based educational research (Knott et al., 2022). An artefact rubric assessed student work along three dimensions: computational accuracy, quality of case-based reasoning, and articulation of the social purpose of the obligation. Instrument validation proceeded in three stages. Three experts, one in Islamic law, one in instructional design and one in qualitative methodology, reviewed all instruments for relevance and clarity; item-level content validity indices ranged from 0.83 to 1.00 with a scale-level index of 0.92, and four items were rewritten following the first round. The instruments were then piloted with a non-participating Grade VIII class in a comparable institution, after which two ambiguous interview prompts were reformulated. Reliability was addressed through inter-rater procedures rather than internal-consistency statistics, which are inappropriate to instruments of this type: two raters independently scored a randomly selected quarter of the artefacts, yielding a weighted kappa of 0.83, and disagreements were resolved through discussion before the remaining artefacts were scored.

Data collection extended across sixteen weeks and generated four corpora. Eight non-participant lesson observations, each of eighty minutes, covered four instructional cycles, with two consecutive lessons observed per cycle; observations were audio-recorded and supplemented by field notes. Semi-structured interviews were conducted with the teacher on three occasions, at the beginning, midpoint and conclusion of the sequence, and once with each institutional informant, with a mean duration of fifty-two minutes. Stimulated-recall interviews of approximately thirty-five minutes were conducted with each focal student after cycles two and four. One focus group discussion with eight students, selected to include focal and non-focal participants, was held at the close of the sequence. Ninety-six student artefacts, comprising four pieces of work from each of the twenty-four students, were collected, together with lesson plans, the institutional curriculum document and the teacher's assessment records. Ethical clearance was obtained from the ethics committee of the sponsoring institution prior to fieldwork. Written permission was secured from the head of the pesantren; written informed consent was obtained from the teacher and institutional informants; and, because all student participants were minors, written parental consent and student assent were obtained separately, with an explicit statement that participation was voluntary, that withdrawal carried no academic consequence, and that no data would inform grading. All personal identifiers were replaced with pseudonyms at the point of transcription, and recordings were stored on encrypted media accessible only to the research team.

Analysis followed reflexive thematic analysis conducted in six recursive stages: familiarisation through repeated reading and re-listening, systematic generation of initial codes, construction of candidate themes, review of themes against the full data set, definition and naming, and integrative writing (Braun & Clarke, 2021). Coding proceeded inductively in the first cycle and was subsequently brought into dialogue with the phase structure of discovery learning and with the fixed-versus-inferable distinction that emerged as analytically central; the resulting scheme is therefore neither purely emergent nor imposed. NVivo 14 supported coding, matrix queries and retrieval, and Microsoft Excel was used to tabulate artefact rubric scores descriptively. Data-source triangulation across observation, interview and artefact corpora was the principal credibility strategy, supported by member checking in which the teacher and four focal students reviewed a summary of preliminary interpretations and confirmed or qualified them, and by peer debriefing with two researchers unconnected to the study. Transferability rests on thick description of setting, participants and instructional sequence sufficient for readers to judge applicability elsewhere. Dependability and confirmability were addressed through a documented audit trail comprising raw materials, coding decisions, analytic memoranda and successive theme structures, and through a reflexivity journal in which the researcher's own formation in Islamic education, a potential source of interpretive bias toward normative readings of classroom events, was recorded and revisited during analysis

III. RESULTS AND DISCUSSION

The study generated 41 hours of recorded material and 96 student artefacts. Table 1 summarises participant characteristics.

Table 1. Participant characteristics and data contribution

Participant group	n	Selection criterion	Data contributed
Fiqh teacher	1	Subject responsibility; ≥ 5 years experience; prior discovery learning training	3 interviews; 8 observed lessons; lesson plans
Whole-class students (Grade VIII)	24	Enrolment in the observed class	96 artefacts; classroom observation
Focal students	6	Maximum variation in prior attainment, residence duration, participation	12 stimulated-recall interviews
Focus group participants	8	Mixed focal and non-focal representation	1 focus group discussion
Institutional informants	2	Curriculum and residential responsibility	2 interviews; policy documents

Reflexive thematic analysis produced four themes comprising eleven sub-themes. Table 2 reports the structure together with the distribution of coded references across data sources, which indicates the evidential base of each theme rather than its importance.

Table 2. Thematic structure and distribution of coded references by source

Theme	Sub-themes	Observation	Interview	Artefact	Total
1. Relocating the object of discovery	Fixed elements taught expositively; inferable elements opened to inquiry; explicit epistemic signalling	58	41	22	121
2. Calibrated autonomy under textual authority	Guidance titrated by phase; the text as data source; teacher as arbiter of last resort	47	39	9	95
3. From computation to disposition	Threshold fluency; case-based adjudication; articulation of social purpose	31	28	61	120
4. Structural friction	Time compression; examination logic; heterogeneous prior knowledge; material scarcity	26	44	7	77

Observation across all four cycles showed a stable and deliberate division of the content. Elements whose status is textually settled, principally the categories of eligible recipients, the threshold quantities and the applicable proportions, were delivered expositively and were never presented as objects of student inquiry. Elements requiring judgement, principally whether a described contemporary situation falls within an established category, were consistently withheld and assigned to students. In the second cycle, for example, the teacher introduced the ratio for agricultural produce directly and in under four minutes, then devoted the remaining fifty-one minutes to a set of six constructed cases in which students had to determine which rule applied and why. The teacher articulated this division explicitly rather than leaving it implicit.

“The threshold is not mine to negotiate and it is not theirs to find. What is theirs is the question of whether this particular income, earned in this particular way, is the thing the threshold speaks about. That question is real, and I refuse to answer it for them”.

Students perceived the distinction and, importantly, did not experience the expositive segments as a contradiction of their agency. One focal student contrasted the two registers directly.

“When Ustadz says the amount, that is the amount, we write it. But when he gives the case of my uncle who sells things online, nobody in the room knows the answer at the start, and that part I remember longest”.

This finding qualifies the constructivist premise that underwrites discovery learning as it is usually formulated. Inquiry research has developed within domains in which the target concept is, in principle, recoverable from evidence, and the design question has therefore been how much support learners require to recover it (Lazonder & Harmsen, 2016; Muhamad Dah et al., 2024). In a normative domain that question is preceded by a prior one: what, in this content, is recoverable at all. The pattern observed here answers it by displacing the object of discovery from the ruling to the application, and in doing so it converts an apparent theological obstacle into a design principle. The result is consistent with evidence that inquiry works best when the task is genuinely open at the point where students are asked to think and closed where openness would merely generate error (Sinha & Kapur, 2021), and it is consistent with cognitive-load reasoning, since asking novices to reconstruct arbitrary conventional quantities would impose extraneous load with no compensating conceptual return (Sweller, 2022). It also extends the pesantren literature, which has documented the persistence of textual authority in these institutions (Kawakip, 2023; Jusubaidi et al., 2024) without specifying how such authority might coexist with learner-centred method. The present case indicates that the two are not in competition once the epistemic boundary is drawn explicitly, and that explicit signalling of that boundary is itself a pedagogical act rather than a concession.

The six discovery phases were enacted with uneven fidelity, and the pattern of unevenness was systematic rather than random. Table 3 summarises enactment across the four cycles.

Table 3. Enactment of discovery phases across four instructional cycles

Phase	Mean minutes	Observed adaptation	Fidelity
Stimulation	7	Replaced abstract prompt with a local case involving a resident family	High
Problem statement	9	Teacher supplied two of three problem formulations in cycles 1–2; students supplied all three by cycle 4	Moderate rising to high
Data collection	18	Classical text excerpts and a translated summary substituted for open searching	Moderate
Data processing	21	Small-group adjudication with a structured worksheet	High
Verification	16	Cross-group challenge; teacher intervened as arbiter in disputed cases	Moderate
Generalisation	9	Compressed or omitted in cycles 1 and 3 under time pressure	Low to moderate

Guidance was not withdrawn as the sequence progressed; it was redistributed. In early cycles the teacher supplied problem formulations and prescribed the sequence of group work; by the fourth cycle students framed their own questions but the teacher intensified intervention at verification, where doctrinal error carried consequences that procedural error did not.

“I can let them argue about whether the debt should be deducted. I cannot let them leave the room believing something incorrect about the obligation itself. So I stand back at the beginning and I come forward at the end”.

The inversion is theoretically informative. Standard fading models assume that guidance recedes monotonically as competence grows, and meta-analytic evidence associates the benefits of inquiry with the presence of guidance without specifying its temporal distribution (Lazonder & Harmsen, 2016). What the present case shows is guidance that is phase-differentiated rather than merely graduated: minimal at problem formulation, moderate during processing, and maximal at verification. This distribution aligns with the problem-solving-then-instruction sequence for which the strongest experimental support exists, in

which exploratory struggle acquires its value from the consolidation that follows it (Sinha & Kapur, 2021), and it converges with findings that fully open inquiry is the least dependable variant precisely because consolidation is weakest there (Muhamad Dah et al., 2024). The finding also contrasts instructively with reports from Islamic religious education elsewhere, where teachers attempting learner-centred method have been observed to relinquish authority wholesale and then reassert it abruptly when discussion produced doctrinally uncomfortable outcomes (Alabdulhadi & Alkandari, 2024). The teacher in this case avoided that oscillation because the boundary described under Theme 1 had already determined where authority would be exercised, which suggests that the two themes are functionally linked rather than merely co-present.

Artefact scores across the four cycles are reported in Table 4. The three rubric dimensions moved at markedly different rates.

Table 4. Mean artefact rubric scores by dimension across four cycles (n = 24; scale 1–4)

Dimension	Cycle 1	Cycle 2	Cycle 3	Cycle 4	Change
Computational accuracy	2.6	2.9	3.0	3.1	+0.5
Case-based reasoning	1.8	2.4	2.9	3.3	+1.5
Articulation of social purpose	1.5	2.0	2.6	3.2	+1.7

Computational accuracy was comparatively high at the outset and improved modestly, which is consistent with the fact that students had encountered thresholds and ratios in prior instruction. The dimensions that shifted substantially were those requiring judgement and justification. By the fourth cycle, artefacts routinely contained conditional reasoning of a kind absent from the first, with students distinguishing between gross and net holdings and between entitlement categories on grounds of need rather than status. The focus group indicated that this shift extended beyond the assessed task.

“Before this I thought zakat was something my father does with a calculator. Now I think it is a question about who in our village is actually poor, which is harder”.

“We argued for twenty minutes about whether a student who receives money from home is a recipient or not. In the end we decided it depends on what the money is for. I had never had to decide anything in fiqh before”.

This finding speaks directly to the concern that animates the zakat literature. Survey research has established that literacy predicts compliance more reliably than awareness (Malik et al., 2024) and that knowledge conditions the translation of religiosity into behaviour (Sadallah et al., 2023), but these studies necessarily treat literacy as a static attribute measured at one point. The present data show literacy being formed, and show that its components form at different rates: the computational component, which surveys most readily capture, was the least responsive to the intervention, while the interpretive and dispositional components, which surveys capture poorly, were the most responsive. The implication is uncomfortable for the measurement conventions of the field, since instruments dominated by threshold and ratio items would have detected only the smallest of the three changes observed here. The finding also converges with meta-analytic evidence that inquiry effects are larger for higher-order outcomes than for factual recall (Antonio & Prudente, 2024) and with evidence that motivational and dispositional gains can precede or exceed measured achievement gains (Toma, 2022). Where the present study departs from that literature is in the moral character of the disposition at issue: what improved was not curiosity about a natural phenomenon but a sense of responsibility toward identifiable others, an outcome for which the science-based inquiry literature offers no direct analogue and which connects instead to work on value formation in religious education (Lafrarchi, 2020; Memon et al., 2023).

Three constraints recurred across the data and are reported here because they bear directly on the transferability of the design. The most consistently reported was time. The generalisation phase, in which students would have articulated a transferable principle from the cases adjudicated, was compressed in two cycles and omitted in one, and the teacher identified this as the principal compromise.

“The syllabus does not have room for the last twenty minutes of what I want to do. Something has to go, and what goes is always the part where they say what it all means”.

The second constraint was assessment. Institutional examinations remained predominantly recall-oriented, creating a divergence between what the method developed and what the system measured, a misalignment that both institutional informants acknowledged. The third was heterogeneity of prior knowledge, unusually pronounced in a residential setting where students arrive at different ages from different educational backgrounds; groups containing students with strong classical text preparation tended to converge quickly on a defensible answer, which shortened deliberation and, in two observed instances, effectively excluded less-prepared members from the reasoning.

These constraints are not idiosyncratic. Comparative evidence from Asian classrooms identifies examination pressure and curricular density as the primary suppressors of active learning approaches, independent of teacher commitment (Ting et al., 2023). Analyses of Indonesian Islamic education policy describe a persistent gap between reform ambition and institutional capacity, in which pedagogical innovation is nominally mandated while assessment and workload structures continue to reward transmission (Kosim et al., 2023). The heterogeneity finding, meanwhile, qualifies an optimistic reading of collaborative inquiry: expertise asymmetry within groups can produce efficient but pedagogically hollow outcomes, a risk that guidance-focused inquiry research anticipates in principle (Lazonder & Harmsen, 2016) but that is intensified in residential institutions where the range of prior preparation is wider than in age-graded day schools (Kawakip, 2023).

Theoretically, the study proposes that discoverability be treated as a property of content rather than a property of method. Existing frameworks specify how much guidance a learner needs; the present analysis indicates that this question is under-determined until one has specified what, within the content, could be discovered at all. Bounded discovery names the resulting design logic: fixed elements are transmitted efficiently and honestly, inferable elements are opened, and the boundary between them is made explicit to learners as part of the instruction. This is a genuine extension rather than a restatement, because it identifies a variable, epistemic status of content, that the guidance literature has held constant by working almost exclusively in domains where all target content is in principle recoverable (Alarcón et al., 2023; Muhamad Dah et al., 2024). The account also has purchase beyond religious education, since law, professional ethics, medical protocol and safety-critical procedure all combine non-negotiable elements with genuinely inferable application, and all have been poorly served by inquiry models that treat every concept as reconstructible.

Practically, the findings support a specific design sequence for teachers of normative content: classify content by epistemic status before planning, transmit the fixed component briefly and without apology, invest the released time in constructed local cases, and concentrate guidance at verification rather than at initiation. The observed rate differential across rubric dimensions further suggests that teachers should assess case reasoning and justification directly rather than inferring them from computational accuracy, since the latter proved the least sensitive indicator of what the instruction actually produced. The use of locally recognisable cases, involving village families, boarding allowances and small online enterprises, appears to have been a material factor in the dispositional shift and is straightforwardly reproducible.

At the policy level, three implications follow. Teacher preparation for religious subjects should address epistemic classification of content explicitly, since the reported difficulty teachers experience with learner-centred method in these subjects (Alabduhadi & Alkandari, 2024; Memon et al., 2023) is at least partly a problem of not knowing what may safely be opened. Assessment reform is a precondition rather than a complement of pedagogical reform, given that recall-dominant examinations systematically undervalue exactly the outcomes the method produces (Kosim et al., 2023). Finally, the alignment between what this instruction developed and what zakat institutions require of future payers (Abdullah & Sapiei, 2018; Sadallah et al., 2023) constitutes a case for treating school-level zakat instruction as an instrument of zakat policy, with corresponding investment in materials and teacher development.

Four limitations qualify these claims. The design is a single case in one institution, and while the analytic ambition is generalisation to theory rather than to a population, the transferability of the enactment pattern to non-residential madrasah or to general schools with weaker textual traditions remains an empirical question. The absence of a comparison condition means that the rubric trajectory reported in Table 4 cannot be attributed to the method with the confidence a controlled design would afford; maturation and the effect of repeated exposure to the assessment format are plausible partial explanations. The observation period, one semester, is too short to establish whether the dispositional gains persist or translate into behaviour, which is precisely the outcome the zakat literature cares about (Malik et al., 2024). Finally, the researcher's own formation in Islamic education, documented reflexively throughout, may have inclined the analysis toward normatively favourable readings of teacher decisions; peer debriefing and member checking mitigate but cannot eliminate this influence.

IV. CONCLUSION

This study set out to describe how discovery learning is enacted in zakat instruction in an Indonesian Islamic boarding school, to characterise the learning it generates, and to identify the conditions that shape it. Three findings answer those questions. The teacher relocated the object of discovery from the ruling to the reasoning that connects a ruling to an unfamiliar case, transmitting doctrinally fixed elements expositively while opening genuinely inferable elements to inquiry, and made this boundary explicit to students. Guidance was phase-differentiated rather than monotonically faded, reaching its maximum at verification, where doctrinal error carried consequences that procedural error did not. Learning gains were unequally distributed across outcome types, with case-based reasoning and articulation of social purpose advancing substantially further than computational accuracy over four instructional cycles, while time compression, recall-oriented assessment and heterogeneous prior preparation constrained the deepest phase of inquiry.

The scientific contribution is the specification of discoverability as a property of content, which supplies inquiry research with a variable it has largely held constant by operating in domains where all target knowledge is in principle reconstructible. Theoretically, bounded discovery reconciles constructivist commitments with revealed-knowledge domains without diluting either, and it does so by relocating rather than restricting learner agency. Practically, the study offers teachers of normative content a defensible planning sequence and a warning that computational assessment will understate what such instruction achieves. For policy, it indicates that teacher preparation should address epistemic classification directly, that assessment reform conditions the viability of pedagogical reform, and that classroom zakat instruction warrants recognition as an instrument of zakat development rather than a purely curricular matter.

Four directions for further work follow. Longitudinal designs are needed to establish whether dispositional gains observed at school age persist into adult compliance behaviour, the link the existing zakat literature assumes but has not traced. Comparative multi-site studies across residential and non-residential institutions would test whether the enactment pattern reported here is a property of the method or of the pesantren setting. Quasi-experimental work with a comparison condition would strengthen causal claims about the rate differential across outcome types. Finally, the transferability of bounded discovery to other normative domains, including legal, medical and professional-ethics education, is an open and consequential question that this single case can raise but not settle.

REFERENCES

- Abdullah, M., & Sapiei, N. S. (2018). Do religiosity, gender and educational background influence zakat compliance? The case of Malaysia. *International Journal of Social Economics*, 45(8), 1250–1264. <https://doi.org/10.1108/IJSE-03-2017-0091>
- Agussuryani, Q., Sudarmin, S., Sumarni, W., Cahyono, E., & Ellianawati, E. (2022). STEM literacy in growing vocational school student HOTS in science learning: A meta-analysis. *International Journal of Evaluation and Research in Education*, 11(1), 51–60. <https://doi.org/10.11591/ijere.v11i1.21647>

- Alabdulhadi, M. M. J., & Alkandari, K. M. (2024). Practices of Islamic education teachers in promoting moderation (wasatiyyah) values among high school students in Kuwait: Challenges and obstacles. *Cogent Education*, 11(1), 2365577. <https://doi.org/10.1080/2331186X.2024.2365577>
- Alarcón, D. A. U., Talavera-Mendoza, F., Paucar, F. H. R., Caceres, K. S. C., & Viza, R. M. (2023). Science and inquiry-based teaching and learning: A systematic review. *Frontiers in Education*, 8, 1170487. <https://doi.org/10.3389/feduc.2023.1170487>
- Antonio, R. P., & Prudente, M. S. (2024). Effectiveness of inquiry-based approaches on students' higher-order thinking skills in science: A meta-analysis. *International Journal of Education in Mathematics, Science and Technology*, 12(1), 251–281. <https://doi.org/10.46328/ijemst.3216>
- Braun, V., & Clarke, V. (2019). Reflecting on reflexive thematic analysis. *Qualitative Research in Sport, Exercise and Health*, 11(4), 589–597. <https://doi.org/10.1080/2159676X.2019.1628806>
- Braun, V., & Clarke, V. (2021). One size fits all? What counts as quality practice in (reflexive) thematic analysis? *Qualitative Research in Psychology*, 18(3), 328–352. <https://doi.org/10.1080/14780887.2020.1769238>
- Jusubaidi, J., Lindgren, T., Mujahidin, A., & Rofiq, A. C. (2024). A model of transformative religious education: Teaching and learning Islam in Pondok Modern Darussalam Gontor, Indonesia. *Millah: Journal of Religious Studies*, 23(1), 171–212. <https://doi.org/10.20885/millah.vol23.iss1.art6>
- Karim, A., Fathurrohman, O., Muhammadun, M., Saripudin, W., Rahmat, D., & Mansir, F. (2023). Altruistic works, religion, and corruption: Kiai's leadership to shape anti-corruption values in pesantren. *Cogent Social Sciences*, 9(1), 2238968. <https://doi.org/10.1080/23311886.2023.2238968>
- Kawakip, A. N. (2023). The practice of shared values and Islamic educational identity: Evidence from a pesantren in East Java, Indonesia. *Journal of Indonesian Islam*, 17(1), 27–53. <https://doi.org/10.15642/JIIS.2023.17.1.27-53>
- Knott, E., Rao, A. H., Summers, K., & Teeger, C. (2022). Interviews in the social sciences. *Nature Reviews Methods Primers*, 2(1), 73. <https://doi.org/10.1038/s43586-022-00150-6>
- Kosim, M., Muqoddam, F., Mubarak, F., & Laila, N. Q. (2023). The dynamics of Islamic education policies in Indonesia. *Cogent Education*, 10(1), 2172930. <https://doi.org/10.1080/2331186X.2023.2172930>
- Lafrarchi, N. (2020). Assessing Islamic religious education curriculum in Flemish public secondary schools. *Religions*, 11(3), 110. <https://doi.org/10.3390/rel11030110>
- Lazonder, A. W., & Harmsen, R. (2016). Meta-analysis of inquiry-based learning: Effects of guidance. *Review of Educational Research*, 86(3), 681–718. <https://doi.org/10.3102/0034654315627366>

- Liu, C., Zowghi, D., Kearney, M., & Bano, M. (2021). Inquiry-based mobile learning in secondary school science education: A systematic review. *Journal of Computer Assisted Learning*, 37(1), 1–23. <https://doi.org/10.1111/jcal.12505>
- Malik, A., Alrasyid, H., & Kamaruddin, M. M. (2024). Student compliance in paying zakat: Do zakat literacy and awareness matter? *Asian Journal of Islamic Management*, 6(1), 63–73. <https://doi.org/10.20885/AJIM.vol6.iss1.art6>
- Marion, Z., Abdullah, A. H., & Rahman, S. N. S. A. (2023). The effectiveness of the GeoGebra-assisted inquiry-discovery learning strategy on students' mastery and interest in algebraic expressions. *International Journal of Information and Education Technology*, 13(11), 1681–1695. <https://doi.org/10.18178/ijiet.2023.13.11.1977>
- Memon, N. A., Alhashmi, M., & Shamma, M. (2023). Islamic schooling and pedagogical transformation in the twenty-first century. *Religions*, 14(9), 1127. <https://doi.org/10.3390/rel14091127>
- Muhamad Dah, N., Mat Noor, M. S. A., Kamarudin, M. Z., & Syed Abdul Azziz, S. S. (2024). The impacts of open inquiry on students' learning in science: A systematic literature review. *Educational Research Review*, 43, 100601. <https://doi.org/10.1016/j.edurev.2024.100601>
- Rohman, F. (2022). Problem based learning in Islamic religious education: The case of the Indonesian pesantren. *Global Journal Al-Thaqafah*, 12(1), 82–97. <https://doi.org/10.7187/GJAT072022-5>
- Sadallah, M., Abdul-Jabbar, H., Bin-Nashwan, S. A., & Abdul Aziz, S. A. (2023). Alms tax (zakat) compliance intention among entrepreneurs from a social cognitive perspective: The moderating role of knowledge. *Journal of Islamic Accounting and Business Research*, 14(8), 1133–1151. <https://doi.org/10.1108/JIABR-04-2022-0104>
- Sinha, T., & Kapur, M. (2021). When problem solving followed by instruction works: Evidence for productive failure. *Review of Educational Research*, 91(5), 761–798. <https://doi.org/10.3102/00346543211019105>
- Sweller, J. (2022). The role of evolutionary psychology in our understanding of human cognition: Consequences for cognitive load theory and instructional procedures. *Educational Psychology Review*, 34(4), 2229–2241. <https://doi.org/10.1007/s10648-021-09647-0>
- Ting, F. S. T., Shroff, R. H., Lam, W. H., Garcia, R. C. C., Chan, C. L., Tsang, W. K., & Ezeamuzie, N. O. (2023). A meta-analysis of studies on the effects of active learning on Asian students' performance in science, technology, engineering and mathematics (STEM) subjects. *The Asia-Pacific Education Researcher*, 32(3), 379–400. <https://doi.org/10.1007/s40299-022-00661-6>

Toma, R. B. (2022). Confirmation and structured inquiry teaching: Does it improve students' achievement motivations in school science? *Canadian Journal of Science, Mathematics and Technology Education*, 22(1), 28–41. <https://doi.org/10.1007/s42330-022-00197-3>