

Cultivating Emotional Awareness in Early Childhood through Mindfulness-Based Play: A Classroom Action Research Reconsidered through Secular and Islamic Contemplative Frameworks

¹Eki Sumiati

Institut Agama Islam Persis Bandung, Indonesia
Email: ekisumiati19@gmail.com

²Himmatul Aliyah Shofia Bahjah Aliansyah

Universitas Islam Negeri Sunan Gunung Djati Bandung, Indonesia
Email: hilyaaliensyah@gmail.com

Correspondence: ekisumiati19@gmail.com

Abstract: Emotional awareness is treated in developmental literature as a load-bearing structure, not a decorative feature, of early social competence. Children who cannot name what they feel do not simply lack a skill; they lack the interpretive apparatus through which every later negotiation with peers, teachers, and their own impulses gets conducted. This study reports a two-cycle classroom action research (Penelitian Tindakan Kelas) conducted with seven children aged 5–6 years at PAUD Al-Muhsinin, testing a sequence of mindfulness-based play activities, including feeling cards, an emotion tree, expressive movement games, and simple breath awareness exercises, as a means of cultivating emotional recognition, verbal labelling, and early empathic response. Unlike much of the applied literature on classroom mindfulness, which tends to report gains as though the underlying construct and its measurement were settled matters, this article treats both as contested. It situates the intervention against two bodies of scholarship that are rarely read together: the social cognitive neuroscience account of mindfulness as a self-regulatory mechanism, including its unresolved evidentiary status, and the Islamic contemplative tradition of *tazkiyat al-nafs* and *tarbiyat al-wijdan*, in which emotional refinement is inseparable from moral and theological formation rather than a value-neutral psychological technique. Qualitative observation across two cycles indicated a shift from reactive, undifferentiated emotional expression toward more differentiated labelling and rudimentary empathic statements, alongside a persistent minority of children for whom the intervention produced compliance rather than internalization, a distinction the action-research design was not built to resolve. The article argues that the pedagogical value of the intervention is real but structurally overstated by the single-group, teacher-as-researcher design that produced it, and it proposes a revised methodological and theoretical agenda, including comparison conditions, standardized instruments, and explicit engagement with Islamic early childhood education (PIAUD) epistemology, for future inquiry in this field.

Keywords: emotional awareness; early childhood; mindfulness; *tazkiyat al-nafs*, classroom action research; social-emotional learning.

Abstrak: Kesadaran emosi merupakan struktur penyangga, bukan pelengkap, dari kompetensi sosial anak usia dini. Artikel ini melaporkan penelitian tindakan kelas dua siklus terhadap tujuh anak usia 5–6 tahun di PAUD Al-Muhsinin yang menguji rangkaian permainan berbasis mindfulness, meliputi kartu perasaan, pohon emosi, permainan gerak ekspresif, dan latihan kesadaran napas, sebagai sarana menumbuhkan pengenalan emosi, pelabelan verbal, dan respons empatik awal. Berbeda dari kecenderungan literatur terapan yang melaporkan hasil positif seolah konstruk dan alat ukurnya sudah mapan, artikel ini memperlakukan keduanya sebagai perkara yang masih diperdebatkan, dengan mendudukan intervensi dalam dua tradisi keilmuan yang jarang dipertemukan, yaitu neurosains kognitif-sosial mindfulness sebagai mekanisme regulasi diri berikut status buktinya yang belum tuntas, dan tradisi kontemplatif Islam *tazkiyat al-nafs* serta *tarbiyat al-wijdan* yang menempatkan penghalusan emosi sebagai bagian tak terpisahkan dari pembentukan moral-teologis, bukan teknik psikologis yang netral nilai. Artikel ini berargumen bahwa nilai pedagogis intervensi tersebut nyata namun secara struktural dilebih-lebihkan oleh desain kelompok tunggal tanpa kontrol dengan guru sebagai peneliti, dan mengusulkan agenda metodologis serta teoretis lanjutan bagi kajian sejenis.

Kata Kunci: kesadaran emosi; anak usia dini; mindfulness; *tazkiyat al-nafs*; penelitian tindakan kelas; pembelajaran sosial-emosional.

Introduction

The claim that early childhood is a formative window for emotional development has achieved something close to consensus status in developmental psychology, which should itself be treated with a measure of suspicion. Consensus claims travel faster than the evidence that produced them, and by the time a proposition reaches teacher-training manuals it has usually shed the qualifications that made it defensible in the first place. Still, the underlying architecture of the claim survives scrutiny. Denham's (2006) synthesis of the school-readiness literature treats social-emotional competence not as an adjunct to cognitive readiness but as a load-bearing precondition for it, on the grounds that a child who cannot regulate frustration or read a peer's distress signal is a child who cannot reliably access the cognitive curriculum built on top of that social floor. Erikson's (1963) account of the third psychosocial stage, initiative versus guilt, extends the claim by giving it a developmental timetable: the child of three to six years is not merely acquiring emotional vocabulary but negotiating a specific crisis of agency, in which the outcome of that negotiation, initiative or guilt as the dominant self-orientation, depends heavily on how the surrounding adults respond to the child's emotionally charged assertions of will.

Classroom practice in Indonesian early childhood education (PAUD) settings has been slow to absorb this architecture into anything resembling deliberate pedagogy. Emotional regulation is routinely treated as a matter of discipline, to be managed through redirection and reprimand, rather than as a competency to be taught through structured activity in the way that letter recognition or number sense is taught. The consequence, observable in the baseline data reported in Section 5 of this article, is a cohort of five- and six-year-olds who reacted to disappointment and conflict with the same undifferentiated repertoire, crying, shoving, or withdrawal, regardless of the specific emotional content of the triggering event. This is not evidence of individual pathology. It is evidence of an instructional gap.

Mindfulness-based intervention has emerged over the past two decades as one candidate for closing that gap, and its adoption in early childhood settings has outpaced the certainty of the evidence supporting it, a mismatch this article treats as its central methodological problem rather than a footnote to be conceded and set aside. Kabat-Zinn's (1994, 2003) formulation of mindfulness as present-moment, non-judgmental attention

was developed for adult clinical populations and only later adapted downward into pediatric and then early-childhood contexts, a genealogy worth stating plainly because each downward adaptation carries assumptions about attentional capacity, metacognitive access, and verbal self-report that become progressively less defensible the younger the population gets. Zelazo and Lyons (2012) offered the most influential developmental account of why the adaptation might nonetheless work, proposing that mindfulness training exercises both a top-down channel, attentional control exercised through repeated practice, and a bottom-up channel, reduced physiological stress reactivity that frees cognitive resources otherwise consumed by dysregulation. That is a coherent mechanistic story. It is also, on the authors' own framing, a hypothesis rather than a settled finding, and the applied literature that followed it has not always preserved that distinction.

Two factions have since formed around the empirical status of that hypothesis and identifying them plainly is more useful than adjudicating between them prematurely. The first, whom this article will call the efficacy faction, is represented by intervention studies reporting measurable gains from structured mindfulness curricula: the Kindness Curriculum trials showing improved executive function and emotion recognition in preschoolers (Flook et al., 2015, as extended by Haines et al., 2023, whose multi-site randomized trial found teacher-rated gains in social-emotional and academic outcomes across 245 children), and the French cluster-randomized protocol combining mindfulness and yoga-based social-emotional learning in socioeconomically disadvantaged pre-kindergarten classrooms (Courbet et al., 2022). The second, the skeptical faction, is represented most comprehensively by Van Dam et al. (2018), whose critical evaluation in *Perspectives on Psychological Science* catalogued a set of recurring methodological failures across the mindfulness research base as a whole: inconsistent operational definitions of the core construct, reliance on self-report instruments whose validity in populations with limited metacognitive capacity is doubtful, weak or absent active control conditions, and a publication ecology that rewards positive findings disproportionately to their statistical robustness. Zelazo and Lyons (2012) themselves, notwithstanding their status as the theoretical anchor of the efficacy faction, explicitly flagged the shortage of adequately controlled developmental trials at the time of their writing, a caveat that subsequent meta-analytic work has only partially resolved; more recent syntheses continue to report genuinely mixed effects of mindfulness-based interventions on children's executive function and self-regulation, with benefits concentrated among children already at elevated risk rather than distributed evenly across a population (Bockmann & Yu, 2022).

This article is written from inside that unresolved dispute rather than above it. It reports a two-cycle classroom action research conducted at PAUD Al-Muhsinin that tested a locally designed sequence of mindfulness-based play activities against the specific outcome of emotional awareness in five- and six-year-old children, and it uses that empirical material as the occasion for a broader argument about how such interventions ought to be studied, interpreted, and, critically, theorized within an Islamic early childhood education context. Three problems with the existing literature and with the initial draft of this study motivate that broader argument. First, the classroom action research (PTK) format in which the underlying study was originally conducted carries structural validity limitations, principally the collapse of the roles of teacher, interventionist, and evaluator into a single person, that are almost never disclosed with the candor the literature on

educational action research recommends (Newton & Burgess, 2008). Second, the theoretical vocabulary used to justify mindfulness practice in Indonesian Islamic early childhood education settings has imported a largely secular, Buddhist-derived contemplative psychology (Kabat-Zinn, 2003) without asking whether the Islamic tradition possesses its own, structurally distinct account of emotional refinement that would license a different pedagogy, a different set of practices, and a different criterion of success. Third, the reporting conventions of applied classroom studies in this field tend toward narrative affirmation of the intervention's success, a rhetorical habit this article deliberately resists by foregrounding disconfirming and ambiguous cases in the results rather than relegating them to a closing paragraph on limitations.

The remainder of the article proceeds as follows. Section 2 develops a dual theoretical framework, first situating emotional awareness and mindfulness within developmental social cognitive neuroscience, then developing tazkiyat al-nafs and tarbiyat al-wijdan as a structurally distinct Islamic contemplative account of emotional formation, and using the comparison to interrogate what a mindfulness-based PAUD curriculum is actually doing when it claims to build emotional awareness. Section 3 reviews the empirical literature on mindfulness-based intervention in early childhood with explicit attention to the tension between the efficacy and skeptical factions identified above. Section 4 details the study's design, including a candid accounting of the validity threats inherent to single-classroom action research. Section 5 reports the results of both cycles, including cases that resisted the intervention's intended effect. Section 6 discusses those results dialectically, weighing genuine pedagogical gain against rival explanations including maturation, novelty effects, and demand characteristics produced by the teacher-researcher's dual role. Section 7 states the study's limitations without euphemism, and Section 8 concludes.

The empirical base for mindfulness-based intervention in early childhood has grown quickly enough that a simple headcount of studies reporting positive effects would overstate the field's maturity. Courbet et al. (2022) registered a genuinely rigorous cluster-randomized trial, sixty-four classrooms and 761 children in a socioeconomically disadvantaged region of France, combining a mindfulness- and yoga-based social-emotional learning curriculum against a wait-list control, with outcome measures spanning teacher-rated questionnaires, blinded standardized observation, and behavioral tasks rather than self-report alone. That design directly addresses several of the Van Dam et al. (2018) critiques, particularly the demand for active comparison conditions and non-self-report measurement, and its results, not yet available at the time the present study was designed but summarized in subsequent publications from the same trial, are the closest thing this literature has to a methodologically adequate answer to the efficacy question. Haines et al. (2023) offer a complementary multi-site randomized trial, sixteen classrooms and 245 children, comparing the Kindness Curriculum against curriculum-as-usual, and found teacher-rated gains across social-emotional, executive function, and academic domains; the authors' own discussion, to their credit, flags that the study was not preregistered and that teacher-report measures cannot rule out rater expectancy effects, a caveat frequently dropped when the study is cited downstream.

Zelazo and Lyons (2012) remain the theoretical anchor for why any of this should work, and Bockmann and Yu's (2022) review of the self-regulation literature confirms a pattern worth taking seriously on its own terms rather than as an inconvenience to be

minimized: effects are real but heterogeneous, concentrated among children with elevated baseline dysregulation, and inconsistent across the specific self-regulatory sub-skills measured. Van Dam et al. (2018) supply the standing methodological indictment against which every study in this literature, including the one reported here, should be read. Their central claims, definitional inconsistency, self-report vulnerability, weak control conditions, and publication bias toward positive findings, are not abstract concerns. They map with uncomfortable precision onto the specific design limitations of classroom action research as a genre, a mapping made explicit in Section 4.

A second and thinner body of literature concerns the reliability of the action-research genre itself as a vehicle for causal claims about pedagogical effectiveness. Newton and Burgess (2008) distinguish three modes of educational action research, emancipatory, practical, and knowledge-generating, and argue that only the third mode makes anything resembling a generalizable causal claim; the first two are legitimately oriented toward local professional improvement and should not be read, or written, as if they were establishing that an intervention works in any transferable sense. The practical implication is that a PTK report structured as a knowledge-generating causal claim, complete with a conclusion recommending the intervention's adoption elsewhere, while actually possessing the evidentiary structure of practical action research, single classroom, no control group, teacher as both implementer and evaluator, is committing a category error regardless of how carefully its narrative is written. The systematic review of action-research methodology in primary science education by Kemmis-derived frameworks (as synthesized in recent systematic literature reviews of the genre) reaches a similar conclusion from a different empirical base: most published classroom action research reports positive outcomes, a pattern more plausibly explained by the genre's structural incentives, teachers do not typically publish accounts of interventions that failed, than by the interventions themselves possessing a uniformly high success rate.

The gap this article addresses sits at the intersection of these two literatures and has, to the authors' knowledge, not been previously articulated in these terms within Indonesian PIAUD scholarship. Existing PTK studies on mindfulness-based emotional development in Indonesian early childhood settings, including the initial draft of the present study, cite the efficacy faction of the mindfulness literature (typically Kabat-Zinn and a handful of downstream applied studies) without engaging the skeptical faction, and they adopt the action-research genre's narrative conventions without engaging the methodological literature on that genre's own limitations. The result is a doubly under-theorized claim: an intervention justified by an incompletely cited psychological literature, evaluated through a research design whose validity constraints are well documented but rarely disclosed, and interpreted without reference to the specific religious and cultural framework, *tarbiyat al-wijdan*, that the institutional setting itself, an Islamic PAUD under a Persis-affiliated foundation, would seem to demand. This article does not claim to resolve either gap definitively. It claims to name both explicitly, to report the present study's findings in a manner that does not paper over the resulting uncertainty, and to propose a revised research agenda in Section 7 capable of closing the gap in subsequent work.

Method

Design and Epistemological Positioning

This study employed classroom action research (Penelitian Tindakan Kelas, PTK) following the Kemmis and McTaggart cyclical model of planning, action, observation, and reflection, conducted across two cycles. The choice of this design was made for the same reason it is usually made in Indonesian PAUD practice, namely that it allows a practicing teacher to investigate and iteratively improve their own classroom practice using resources and access that a purely external, quasi-experimental design would not permit. That justification is legitimate on its own terms. It is not, however, a justification for treating the resulting findings as evidence of the intervention's transferable efficacy, a distinction this article maintains throughout rather than blurring in the manner Newton and Burgess (2008) warn against.

Following their taxonomy, this study is best classified as practical action research oriented toward local pedagogical improvement, not knowledge-generating action research oriented toward causal generalization. The single most consequential validity threat in this classification is the collapse of three normally separated roles, program designer, program implementer, and outcome evaluator, into the person of the classroom teacher-researcher. This collapse introduces at least three specific and non-hypothetical biases into the data reported in Section 5: confirmation bias in the interpretation of ambiguous child behavior during observation, since the observer has a professional and personal stake in the intervention's apparent success; demand characteristics in the children's behavior itself, since young children are acutely sensitive to adult approval and may perform emotional articulation they do not fully possess in order to secure the teacher's evident pleasure; and the absence of any comparison condition capable of distinguishing intervention effects from simple maturation across the roughly eight-week span of the two cycles, a period during which five- and six-year-olds develop substantially regardless of classroom input. None of these threats were eliminated in the present study. They were mitigated, to the extent the design permitted, through the triangulation and reflexive procedures described in Section 4.5, and they are reported here rather than in a closing disclaimer precisely because a reader evaluating the results in Section 5 needs this information before, not after, encountering the findings.

Participants and Setting

The study was conducted at PAUD Al-Muhsinin with a class of seven children in the five-to-six-year-old (Kelompok A) age band, during the even semester of the 2024–2025 academic year. The small cohort size, a structural feature of the single-classroom setting rather than a sampling choice, has direct consequences for the interpretability of any pattern reported below; with $n = 7$, individual children's idiosyncratic responses carry disproportionate weight in any aggregate description, and no claim in this article should be read as generalizable beyond this specific classroom without independent replication. Consent and assent procedures followed the institution's standard practice for classroom-embedded pedagogical activity, and no child was excluded from participation in the activities themselves, since they were integrated into the regular instructional schedule rather than administered as a separate research protocol; observational data collection, by

contrast, was disclosed to parents in advance in accordance with standard research-ethics practice for classroom-based studies involving minors.

Procedure

The intervention consisted of two instructional cycles, each following the plan-act-observe-reflect structure, separated by a reflection period in which the first cycle's observational data directly informed the redesign of the second. Cycle I focused on the foundational task of emotion recognition and labelling across five basic categories, happiness, sadness, anger, fear, and surprise, delivered through illustrated feeling cards depicting facial expressions and a circle-based ball-catching game in which the child receiving the ball named an emotion and narrated a personal episode of having felt it. A guided breath-awareness exercise, framed for the children as "blowing up a balloon" with the hand placed on the abdomen to feel it expand and contract, was introduced as a regulatory anchor preceding the more verbally demanding labelling activities, on the theoretical grounds, following Zelazo and Lyons (2012), that reduced physiological arousal should widen the cognitive bandwidth available for the labelling task that followed it.

Cycle II, redesigned in response to the reflection recorded at the end of Cycle I (detailed in Section 5.2), shifted emphasis from bare labelling toward the coupling of emotional recognition to personal narrative and peer-directed empathic response, the coupling identified in Section 2.3 as the specific signature of a *tarbiyat al-wijdan* rather than a purely secular mindfulness logic. The principal new activity was an "emotion tree," a collaborative visual in which each child attached a leaf bearing an emotion they recalled experiencing and narrated the accompanying episode to the group, followed by a structured opportunity for peers to respond. A parallel drawing activity allowed children uncomfortable with verbal disclosure, a subgroup identified explicitly in the Cycle I reflection, an alternative expressive channel.

Instruments and Data Analysis

Data were collected through structured teacher observation using a checklist derived from the three-part definition of emotional awareness given in Section 2.1 (detection, labelling, contextually appropriate modulation), supplemented by field notes recording verbatim child utterances during the labelling and narrative activities, and by simple frequency counts of children able to correctly identify and verbally label a target emotion within each session. This instrumentation is a deliberate downgrade from the vaguer, purely impressionistic narrative reporting common in the PTK genre, but it remains, by the standards Van Dam et al. (2018) apply to the adult mindfulness literature, methodologically modest: it relies entirely on a single, non-blinded observer who is simultaneously the intervention's implementer, it possesses no standardized psychometric validation, and it generates frequency counts too small to support any inferential statistical procedure. Data were analyzed qualitatively through iterative coding of field notes against the three-part construct, with disconfirming cases, children whose behavior did not track the intervention's intended trajectory, deliberately retained and reported in Section 5 rather than treated as noise to be averaged away.

Trustworthiness and Ethical Considerations

Three procedures were used to partially offset the single-observer limitation described in Section 4.1, without claiming to eliminate it. First, field notes were reviewed jointly by both authors at the end of each cycle, providing a minimal form of investigator triangulation. Second, the coding scheme was applied to verbatim quotations rather than to the observer's summary impressions wherever possible, reducing, though not removing, the space for motivated interpretation. Third, the Cycle I reflection explicitly recorded the subgroup of children for whom the intervention appeared to produce compliance rather than internalized change, a category most researchers working within the efficacy-affirming conventions of the PTK genre would have been tempted to omit, and that omission is treated here as itself a threat to trustworthiness worth naming. Ethical procedures followed standard protections for research involving young children, including parental notification, the option to decline observational data collection without exclusion from the underlying instructional activity, and de-identification of all child data reported in Section 5.

Results and Discussion

Baseline Observation

Prior to intervention, baseline observation recorded a cohort in which emotional labelling capacity was low and unevenly distributed. When the teacher read a story with explicit emotional content, only one of ten children present at that baseline session (a slightly larger group than the eventual seven-child intervention cohort, reflecting normal attendance variation) could answer content questions tied to the characters' emotional states; the remainder displayed poor sustained attention, difficulty remaining seated, and, in several documented instances, reactive behavior, crying, shoving a peer during a minor conflict, or withdrawing physically when disappointed, that gave no indication the underlying emotional state had been consciously registered by the child at all. This baseline is reported in some detail because it establishes the floor against which the Cycle I and Cycle II gains described below must be read; a cohort starting this low has considerable room for improvement attributable to simple maturation and increased classroom exposure, independent of any specific intervention effect, a point developed further in Section 6.2.

Cycle I: Emotion Recognition

Table 1 summarizes the distribution of observed labelling accuracy among the seven children who completed Cycle I's circle-based ball-catching activity.

Table 1 Distribution of Emotion-Labelling Accuracy, Cycle I (n = 7)

Category	n	Description
Accurate and spontaneous labelling	3	Child correctly named the target emotion and volunteered a related personal episode without prompting
Partial or confused labelling	2	Child confused sadness and anger, or required repeated prompting to produce a label

Category	n	Description
Low engagement	2	Child participated only imitatively, following peers' actions without independent verbal labelling

The expressive movement activity that accompanied the labelling task, jumping while smiling for happiness, sitting with a lowered head for sadness, clenching fists while exhaling forcefully for anger, produced markedly more consistent participation across the full cohort than the verbal labelling task did on its own, a divergence the Cycle I reflection interpreted as evidence that embodied, non-verbal channels of emotional expression were developmentally more accessible to this cohort than verbal self-report, a finding directly relevant to the instrument limitations discussed in Section 4.4, since any evaluation relying primarily on verbal self-report would have systematically underestimated the emotional awareness of children who could express, but not yet narrate, their internal states.

The Cycle I reflection identified four specific limitations requiring redesign: the media used were judged insufficiently concrete for the two low-engagement children, an initial “hook” activity such as a song or puppet was judged necessary to secure attention before the more cognitively demanding labelling task began, an individualized approach was judged necessary for passive children rather than relying on whole-group activities alone, and more time was judged necessary for reflective discussion than the original session structure allowed. These four limitations directly informed the Cycle II redesign described in Section 4.3.

Cycle II: Narrative Coupling and Emergent Empathy

Cycle II’s emotion-tree activity produced the clearest qualitative shift observed across the study. Children who had been coded as low-engagement in Cycle I participated in the narrative-sharing component of the emotion tree, several offering unprompted personal disclosures (“I felt sad when my toy broke”) that had not appeared in the more structured Cycle I format, a shift consistent with the reflection’s hypothesis that embodied and visual channels, rather than purely verbal prompting, were the more developmentally appropriate entry point for this cohort.

Table 2 summarizes the qualitative shift between cycles across the three sub-capacities defined in Section 2.1.

Table 2 Qualitative Shift in Emotional Awareness Sub-Capacities, Cycle I to Cycle II

Sub-capacity	Cycle I	Cycle II
Detection	Inconsistent; reactive behavior without visible recognition	More consistent self-report of internal states preceding action
Labelling	Accurate for 3/7 children; confused for 2/7; absent for 2/7	Expanded to the majority of the cohort, with two children still requiring prompting

Sub-capacity	Cycle I	Cycle II
Modulation and prosocial coupling	Not observed	Emergent in multiple children, including spontaneous empathic statements (“I can help if you’re sad”) and self-reported use of the breath-awareness strategy prior to reacting to frustration

The emergence of spontaneous empathic statements in Cycle II is treated in this article as the single most theoretically significant finding, since it is precisely the coupling of emotional labelling to a morally patterned response, rather than bare labelling alone, that the tarbiyat al-wijdan framework developed in Section 2.3 predicts and that a purely secular attentional-training account does not straightforwardly anticipate. One child’s reported use of a breath-awareness strategy prior to an anger-triggering event (“if I’m angry, I take a breath first now”) is consistent with the self-regulatory mechanism Zelazo and Lyons (2012) propose, and it is reported here as a single, non-generalizable instance rather than as evidence of a cohort-wide regulatory shift, in keeping with the small-sample caution stated in Section 4.2.

Not every child in the cohort tracked this trajectory, a fact this article reports because the PTK genre’s narrative conventions would ordinarily allow it to be quietly omitted. Two children, coded as low-engagement in Cycle I, continued to participate in Cycle II activities in a manner that field notes describe as imitative rather than independently generated, following the verbal and gestural cues of higher-participating peers without producing the spontaneous, unprompted disclosures that characterized the cohort’s more visible gains. Whether this reflects a genuinely slower developmental trajectory, an underlying temperamental disposition toward reticence that the intervention was not designed to address, or simply a measurement artifact produced by an observation instrument that privileges verbal disclosure even after the Cycle I redesign attempted to correct for that bias, cannot be determined from the present data, and this article declines to resolve the ambiguity in the intervention’s favor.

Interpreting the Gains

The shift documented between Cycle I and Cycle II, from largely reactive, undifferentiated emotional expression toward more differentiated labelling and emergent empathic coupling, is consistent with the pedagogical hypothesis stated in Section 2.1, that structured, play-mediated practice in detecting, labelling, and modulating emotion produces measurable movement toward the readiness outcomes Denham (2006) describes and the psychosocial trajectory Erikson (1963) describes. It is also consistent, at the mechanistic level, with the dual-pathway model Zelazo and Lyons (2012) propose, insofar as the single documented instance of a child deploying breath-awareness prior to an anger-triggering event maps onto exactly the bottom-up, stress-reduction pathway their model specifies. None of this should be read as an endorsement of the intervention’s efficacy in the strong, generalizable sense that a randomized trial such as Courbet et al.’s (2022) is designed to establish. It should be read as evidence that a locally designed, culturally embedded activity sequence produced observable, theoretically coherent change in a

specific classroom, which is precisely the kind of claim practical action research, correctly classified per Newton and Burgess (2008), is equipped to support and no more than that.

The Rival Hypotheses

Three rival explanations deserve explicit consideration before the observed gains are attributed to the intervention itself, and the fact that the original draft of this study did not consider any of them is itself a symptom of the genre problem identified in Section 3. The first rival is simple maturation. Five- and six-year-olds develop substantial emotional and linguistic capacity over an eight-week span under any classroom conditions, and the baseline-to-Cycle-II comparison reported in Section 5 cannot, on the design's own terms, distinguish maturation from intervention effect, since no comparison classroom received an alternative or absent treatment against which the intervention group's trajectory could be benchmarked. The second rival is demand characteristics arising from the collapsed teacher-researcher role identified in Section 4.1. Young children are highly attuned to adult approval, and a teacher who has visibly designed and is enthusiastically delivering an emotion-focused curriculum may elicit performed emotional articulacy, children producing the labels and empathic statements they have learned please the teacher, that does not reflect a stable underlying change in emotional awareness as defined in Section 2.1. The third rival, related but distinct, is a halo effect in the observational data itself: an observer who has invested professional effort in designing an intervention has a documented tendency, independent of conscious bias, to interpret ambiguous child behavior in the direction that confirms the intervention's success, a pattern Van Dam et al. (2018) flag explicitly as endemic to the broader mindfulness literature and that this study's single-observer design does nothing structural to prevent.

None of these rival hypotheses can be ruled out by the data reported in Section 5. They are also not, on balance, sufficient to explain away the full pattern observed, particularly the qualitative texture of the Cycle II narrative disclosures, which included specific, idiosyncratic personal content ("I felt sad when my toy broke") difficult to attribute purely to rehearsed performance, and the persistence of a genuinely disconfirming subgroup, the two imitative-participation children, whose continued resistance to the intervention's intended trajectory is not what a pure demand-characteristics account would predict, since a child purely performing for teacher approval has every incentive to fake the gain the two low-engagement children conspicuously did not display. The honest conclusion is not that the intervention produced no effect, nor that it produced the full effect its narrative description in Section 5 might suggest to an uncritical reader, but that the study's design cannot separate a real, moderate pedagogical effect from a real, moderate measurement artifact, and that the two most likely coexist in whatever gain actually occurred.

Islamic Contemplative Pedagogy as Interpretive Frame

The theoretical argument developed in Section 2.3 receives its strongest empirical support from precisely the finding the secular mindfulness literature has the most difficulty explaining, namely the emergence of spontaneous prosocial and empathic statements coupled to emotional labelling rather than labelling alone. A purely Kabat-Zinn-derived clinical mindfulness curriculum, oriented toward value-neutral present-moment

observation, generates no strong prediction that labelling should couple to moral response in this way; the non-judging facet of secular mindfulness explicitly brackets evaluative and moral content rather than embedding it. What this study's Cycle II activities actually did, whether by original design or by the teacher's implicit cultural formation, was closer to *muraqabah* and *tazkiyat al-nafs* as theorized in Sadiq and Ahmad's (2026) Culturally Embedded Mindfulness Model, in which self-observation is inseparable from moral accountability and in which *niyyah*, the intention behind an emotional or behavioral response, is treated as a precondition for that response's proper formation, not an afterthought to it. The teacher's consistent framing of the emotion-tree activity around gratitude, patience, and mutual care, rather than neutral observation of internal states, is the pedagogically operative variable this article proposes for explicit testing in future, better-controlled research, since it is a variable the existing secular mindfulness literature has no vocabulary to isolate.

This reframing carries a practical implication for PAUD curriculum design that goes beyond the present study's modest evidentiary base. If *tarbiyat al-wijdan*, rather than imported secular mindfulness, is the more accurate description of what already works in Islamic early childhood settings, then future curriculum development in this field should stop importing Kabat-Zinn-derived materials wholesale and translating them into Indonesian, and should instead build activity sequences explicitly grounded in Islamic conceptual vocabulary, *sabr*, *shukr*, *rahmah*, *niyyah*, from the outset, with secular developmental psychology retained as a source of mechanistic hypotheses about attention and stress physiology rather than as the primary source of pedagogical justification. This is not a call to discard Zelazo and Lyons (2012) or the broader contemplative science literature; their attentional and physiological mechanisms plausibly operate regardless of the moral framework in which the practice is embedded. It is a call to stop treating that mechanistic literature as though it were also a sufficient theory of what emotional formation is *for*, a question the Islamic tradition has an answer to that the secular literature, by design, brackets.

Conclusion

Emotional awareness functions as a structural precondition for early social competence, not a peripheral enrichment activity, and the baseline data reported in this study confirm that a substantial proportion of the children observed entered the intervention period without the basic detection and labelling capacities that competence requires. The two-cycle intervention produced a documented, theoretically coherent shift toward more differentiated emotional labelling and, in its second cycle, toward the emergence of morally coupled prosocial response, a pattern more consistent with an Islamic *tarbiyat al-wijdan* framing than with the value-neutral secular mindfulness framing under which the intervention was originally conceived. That shift occurred alongside a persistent subgroup for whom the intervention produced imitation rather than internalization, a finding this article has reported rather than absorbed into an unqualified success narrative. The single-classroom, teacher-as-researcher design that generated these findings carries validity limitations serious enough that the results should be read as a documented local pedagogical improvement, in the practical action-research sense Newton and Burgess

(2008) define, rather than as evidence of the intervention's transferable efficacy. The field's task going forward is not to accumulate further single-classroom narratives affirming that mindfulness-based play helps young children recognize their feelings, a claim already adequately supported at the level of plausibility, but to build the comparison-group, better-instrumented, and theoretically explicit studies, grounded openly in Islamic contemplative pedagogy where the institutional setting warrants it, capable of establishing what, specifically, that help consists of, for whom, and why.

References

- Bockmann, J. O., & Yu, S. Y. (2022). Using mindfulness-based interventions to support self-regulation in young children: A review of the literature. *Early Childhood Education Journal*. <https://doi.org/10.1007/s10643-022-01333-2>
- Courbet, O., Daviot, Q., Kalamarides, V., Habib, M., Castillo, M.-C., & Villemonteix, T. (2022). Promoting psychological well-being in preschool children: Study protocol for a randomized controlled trial of a mindfulness- and yoga-based socio-emotional learning intervention. *Trials*, 23, 1050. <https://doi.org/10.1186/s13063-022-06979-2>
- Denham, S. A. (2006). Social-emotional competence as support for school readiness: What is it and how do we assess it? *Early Education and Development*, 17(1), 57–89. https://doi.org/10.1207/s15566935eed1701_4
- Erikson, E. H. (1963). *Childhood and society* (2nd ed.). W. W. Norton & Company.
- Flook, L., Goldberg, S. B., Pinger, L., & Davidson, R. J. (2015). Promoting prosocial behavior and self-regulatory skills in preschool children through a mindfulness-based kindness curriculum. *Developmental Psychology*, 51(1), 44–51. <https://doi.org/10.1037/a0038256>
- Goleman, D. (1995). *Emotional intelligence: Why it can matter more than IQ*. Bantam Books.
- Haines, B., Hong, P. Y., Immel, K. R., & Lishner, D. A. (2023). The mindfulness-based kindness curriculum for preschoolers: An applied multi-site randomized control trial. *Mindfulness*, 14(9), 2195–2210. <https://doi.org/10.1007/s12671-023-02210-8>
- Kabat-Zinn, J. (1994). *Wherever you go, there you are: Mindfulness meditation in everyday life*. Hyperion.
- Kabat-Zinn, J. (2003). Mindfulness-based interventions in context: Past, present, and future. *Clinical Psychology: Science and Practice*, 10(2), 144–156. <https://doi.org/10.1093/clipsy.bpg016>
- Locke, E. A. (2005). Why emotional intelligence is an invalid concept. *Journal of Organizational Behavior*, 26(4), 425–431. <https://doi.org/10.1002/job.318>
- Newton, P., & Burgess, D. (2008). Exploring types of educational action research: Implications for research validity. *International Journal of Qualitative Methods*, 7(4), 18–30. <https://doi.org/10.1177/160940690800700402>

Sadiq, & Ahmad, M. S. (2026). Islamic mindfulness as cultural mindfulness: A conceptual framework for decision-making and well-being. *Frontiers in Psychology*, 17, Article 1715750. <https://doi.org/10.3389/fpsyg.2026.1715750>

Van Dam, N. T., van Vugt, M. K., Vago, D. R., Schmalzl, L., Saron, C. D., Olendzki, A., Meissner, T., Lazar, S. W., Kerr, C. E., Gorchov, J., Fox, K. C. R., Field, B. A., Britton, W. B., Brefczynski-Lewis, J. A., & Meyer, D. E. (2018). Mind the hype: A critical evaluation and prescriptive agenda for research on mindfulness and meditation. *Perspectives on Psychological Science*, 13(1), 36–61. <https://doi.org/10.1177/1745691617709589>

Zelazo, P. D., & Lyons, K. E. (2012). The potential benefits of mindfulness training in early childhood: A developmental social cognitive neuroscience perspective. *Child Development Perspectives*, 6(2), 154–160. <https://doi.org/10.1111/j.1750-8606.2012.00241.x>