

**ANALISIS RASIONAL INSTRUMEN DALAM PENGEMBANGAN SISTEM
EVALUASI DI TK IT AL IFFAH BEKASI**

Aridatul Kholidah

Sekolah Tinggi Agama Islam Al-Barokah

aridatul@gmail.com

Noviyanti

Sekolah Tinggi Agama Islam Al-Barokah

novi.yanti@almuslim.sch.id

Santy Ika

Sekolah Tinggi Agama Islam

santyika06@gmail.com

Eva Siti Faridah

Sekolah Tinggi Agama Islam Al-Barokah

evasitifaridah86@gmail.com

Abstrak

Evaluasi pembelajaran itu penting dalam pendidikan anak usia dini untuk memantau perkembangan kognitif, afektif, psikomotorik, social, emosional dan motorik anak menggunakan alat ukur yang sesuai dengan tujuan yang diharapkan. Dengan proses evaluasi, guru mendapatkan gambaran keberhasilan siswa. Penelitian ini bertujuan untuk menganalisis rasionalitas instrumen evaluasi yang digunakan di TK IT Al Iffah dalam menilai perkembangan peserta didik. Analisis rasional instrumen diperlukan untuk memastikan bahwa alat ukur yang digunakan sesuai dengan tujuan pembelajaran, indikator perkembangan anak, dan prinsip evaluasi pendidikan anak usia dini. Metode penelitian yang digunakan adalah kualitatif dengan metode pengumpulan data menggunakan observasi, wawancara dan dokumen. Berdasarkan hasil perhitungan dari data Excel (30 responden anak TK IT Al Iffah), diperoleh bahwa rata-rata skor tertinggi terdapat pada **aspek motorik (7,44)**. Hal ini menunjukkan bahwa perkembangan motorik anak lebih menonjol dibanding aspek spiritual. Distribusi skor perilaku disiplin dan social anak menunjukkan bahwa **mayoritas anak (50%) berada pada kategori “baik”**, sementara 23% anak masih memerlukan bimbingan. Nilai **Cronbach’s Alpha sebesar 0,273** menandakan rendahnya konsistensi antar-aspek penilaian. Hal ini bukan semata menunjukkan kelemahan instrumen, tetapi lebih menggambarkan bahwa lima domain perkembangan anak usia dini (kognitif, bahasa, motorik, sosial, spiritual) bersifat **multidimensional**. Hal ini menyebabkan beberapa instrumen tidak memiliki keterkaitan langsung dengan tujuan pembelajaran yang direncanakan. Hal ini menunjukkan bahwa sebagian besar instrumen yang digunakan masih bersifat umum dan belum sepenuhnya mengacu pada aspek validitas isi dan konstruk. Guru telah berupaya melakukan penyesuaian berdasarkan karakteristik anak dan capaian pembelajaran. Penyusunan instrument sangat

penting disusun secara rasional agar hasil evaluasi dapat menggambarkan perkembangan anak secara objektif.

Kata kunci : Analisis Rasional, Instrumen Evaluasi, Pendidikan Anak Usia Dini

Abstract

Learning evaluation is essential in early childhood education to monitor children's cognitive, affective, psychomotor, social, emotional, and motor development using measurement tools aligned with expected goals. Through the evaluation process, teachers gain an overview of student success. This study aims to analyze the rationality of the evaluation instruments used at TK IT Al Iffah in assessing student development. A rational analysis of instruments is necessary to ensure that the measurement tools used align with learning objectives, child development indicators, and the principles of early childhood education evaluation. The research method used is qualitative, with data collection methods including observation, interviews, and documentation. Based on calculations from Excel data (30 respondents from TK IT Al Iffah), the results show that the highest average score is in the motor aspect (7.44). This indicates that children's motor development is more prominent than the spiritual aspect. The distribution of children's disciplined and social behavior scores shows that the majority (50%) are in the "good" category, while 23% still require guidance. The Cronbach's Alpha value of 0.273 indicates low consistency between assessment aspects. This does not merely indicate a weakness in the instrument but rather reflects that the five domains of early childhood development (cognitive, language, motor, social, spiritual) are multidimensional. Consequently, some instruments do not have a direct link to the planned learning objectives. This suggests that most of the instruments used are still general in nature and do not yet fully refer to aspects of content and construct validity. Teachers have made efforts to make adjustments based on child characteristics and learning achievements. It is crucial to develop instruments rationally so that evaluation results can objectively reflect child development.

Keywords : Rational Analysis, Evaluation Instruments, Early Childhood Education

PENDAHULUAN

Evaluasi pembelajaran itu penting dalam sistem pendidikan. Fungsinya bukan hanya untuk mengetahui seberapa jauh proses belajar berjalan, tetapi sebagai alat ukur: sudah sesuai atau belum dengan tujuan yang diharapkan. Dengan proses evaluasi, guru mendapatkan gambaran keberhasilan siswa. Guru dapat mengetahui apa yang kurang, lalu memperbaiki cara mengajar atau strateginya supaya ke depannya lebih efektif. Arikunto menatakan evaluasi itu proses sistematis untuk menilai sudah tercapai atau belum tujuan pembelajaran dengan menggunakan instrumen yang memang terbukti valid dan reliabel¹. Keberhasilan pendidikan sebenarnya sangat tergantung pada seberapa baik alat evaluasinya—harus mempunyai dasar rasional dan ilmiah, tidak asal-asalan.

Pendidikan anak usia dini (PAUD) sendiri adalah masa yang krusial dalam hidup manusia. Di sinilah anak-anak berkembang pesat—tidak hanya dari sisi kognitif, tapi juga

¹ Suharsimi Arikunto, Supardi, and Suhardjono, *Penelitian Tindakan Kelas: Edisi Revisi* (Jakarta: Bumi Aksara, 2015).

emosional, sosial, motorik, sampai perilaku mereka di masa depan ikut terbentuk². Belakangan ini, teknologi pendidikan seperti kecerdasan buatan (AI) makin banyak dipakai di PAUD. Teknologi ini sangat membantu, mulai dari memperkuat proses belajar, sampai memudahkan pemantauan perkembangan anak³.

Evaluasi pendidikan pada PAUD memiliki peran yang sangat vital karena banyak aspek yang harus dinilai bukan hanya akademik—ada kognitif, afektif, sosial-emosional, motorik, spiritual, semuanya harus diperhatikan. Pendidikan anak usia dini memang tidak hanya soal nilai, tetapi tentang membentuk karakter, perilaku, dan kemampuan dasar sesuai tahap perkembangannya. Itu sebabnya, alat evaluasinya pun tidak bisa sembarangan. Harus disusun dengan prinsip perkembangan anak—holistik, kontekstual, dan tetap menyenangkan⁴.

Sayangnya, kenyataan di lapangan masih banyak guru yang belum paham cara membuat alat evaluasi yang benar. Banyak PAUD masih memakai instrumen yang seadanya, kadang hanya meniru dari tahun ke tahun tanpa ada pembaruan atau uji validitas⁵. Akibatnya, hasil evaluasinya sering tidak akurat, bahkan bisa menyesatkan soal kemampuan anak. Di banyak TK, praktik evaluasi juga masih banyak hambatan. Pertama, metode lama terlalu mengandalkan observasi langsung guru, padahal metode ini sering tidak bisa menangkap seluruh perkembangan anak—apalagi kalau anak sedang bermain bebas, sering ada sisi yang terlewat, akhirnya umpan balik kepada guru menjadi terlambat atau kurang mendalam⁶. Kedua, rasio guru dan murid yang tidak seimbang membuat guru kesulitan memantau setiap anak secara real-time, padahal itu sangat penting untuk pengalaman belajar yang bermakna⁷. Ketiga, ada jarak antara aturan atau kurikulum ideal dengan pelaksanaan di lapangan. Pada akhirnya,

² Supriya Khadka and Sanchari Das, “XR Design Framework for Early Childhood Education,” arXiv:2601.18979, preprint, arXiv, January 26, 2026, <https://doi.org/10.48550/arXiv.2601.18979>.

³ Yi Honghu, Liu Ting, and Lan Gongjin, “The Key Artificial Intelligence Technologies in Early Childhood Education: A Review,” arXiv:2401.05403, preprint, arXiv, December 20, 2023, <https://doi.org/10.48550/arXiv.2401.05403>.

⁴ Mulyasa, *Manajemen Paud* (Remaja Rosdakarya, 2014), <https://openlibrary.telkomuniversity.ac.id/pustaka/18540/manajemen-paud.html>.

⁵ Sukardi Sukardi, *Evaluasi Pendidikan: Prinsip & Operasionalnya* (Jakarta: Bumi Aksara, 2017), <https://dpk.kepriprov.go.id/opac/detail/bv10f>.

⁶ Yuanyuan Yang et al., *Validating the Effectiveness of a Large Language Model-Based Approach for Identifying Children’s Development across Various Free Play Settings in Kindergarten*, version 1, 2025, <https://doi.org/10.48550/ARXIV.2505.03369>.

⁷ Sindhuja Penchala et al., “Learning in Focus: Detecting Behavioral and Collaborative Engagement Using Vision Transformers,” version 2, preprint, arXiv, 2025, <https://doi.org/10.48550/ARXIV.2508.15782>.

layanan evaluasi menjadi tidak merata, bahkan bisa ada kesenjangan dalam kualitas maupun inklusivitasnya⁸.

Gronlund & Linn menyatakan bahwa untuk membuat instrumen evaluasi yang benar, ada beberapa syarat penting yang harus dipenuhi: validitas (alat ukurnya memang benar-benar mengukur hal yang seharusnya diukur), reliabilitas (hasilnya konsisten), objektivitas, dan kepraktisan⁹. Jika kriteria ini tidak dipenuhi, hasil penilaian menjadi tidak dapat diandalkan untuk mengambil keputusan di dunia pendidikan.

Dalam pendidikan Islam, evaluasi tidak hanya soal nilai kognitif — lebih dari itu, juga menilai aspek spiritual dan moral. Evaluasi jadi alat muhasabah, ajang introspeksi, buat tahu sejauh mana siswa benar-benar paham, menghayati, dan menjalani nilai-nilai Islam sehari-hari. Al-Qur'an sendiri sudah menekankan pentingnya evaluasi diri, seperti di Surah Al-Hasyr ayat 18:

“Hai orang-orang yang beriman, bertakwalah kepada Allah dan hendaklah setiap diri memperhatikan apa yang telah diperbuatnya untuk hari esok.” (QS. Al-Hasyr: 18)

Ayat ini jelas mengingatkan pentingnya menilai perilaku dan amal sendiri. Dalam pendidikan, ini artinya evaluasi harus menjadi alat untuk memperbaiki proses belajar, supaya hasilnya lebih baik lagi.

Analisis rasional instrumen menjadi langkah penting dalam memastikan bahwa alat ukur yang digunakan benar-benar sesuai dengan tujuan pembelajaran. Rasionalitas dalam konteks ini berarti bahwa setiap komponen instrumen mulai dari indikator, item, rubrik, hingga cara penskoran harus didasarkan pada teori pengukuran dan prinsip evaluasi pendidikan. Melalui analisis rasional, guru dapat menilai apakah setiap item instrumen telah relevan, proporsional, dan dapat diobservasi secara objektif. Evaluasi yang tidak rasional dapat menimbulkan kesalahan dalam penilaian perkembangan anak, baik berupa *underestimate* (menilai terlalu rendah) maupun *overestimate* (menilai terlalu tinggi). Akibatnya, guru tidak dapat memberikan intervensi yang tepat sesuai kebutuhan anak. Dengan demikian, analisis rasional instrumen merupakan upaya untuk meningkatkan validitas dan reliabilitas hasil evaluasi serta menjamin akuntabilitas proses pembelajaran.

⁸ Angela Andreella et al., “A Two-Step Model to Study the Inclusivity’s Distribution of Italian Early Childhood Education and Care Services,” version 1, preprint, arXiv, 2024, <https://doi.org/10.48550/ARXIV.2407.15610>.

⁹ Norman Edward Gronlund and Robert L. Linn, *Measurement and Evaluation in Teaching* (Macmillan, 1990).

TK IT Al Iffah sendiri adalah sekolah Islam terpadu yang mengusung pembelajaran tematik integratif, semuanya berlandaskan nilai-nilai Islam. Visi mereka juga jelas: ingin membentuk generasi yang shalih, cerdas, dan punya karakter Qur’ani. Dalam proses belajar, guru-gurunya pakai macam-macam evaluasi—mulai dari observasi perilaku, portofolio karya anak, sampai lembar penilaian harian. Tapi, dari pengamatan awal, instrumen evaluasi yang dipakai kebanyakan belum punya dasar penyusunan yang benar-benar rasional. Kebanyakan guru membuat instrumennya hanya berdasarkan pengalaman pribadi, belum mengacu ke prinsip ilmiah pengukuran. Kadang indikator perkembangannya terlalu umum, kayak “anak mampu bekerja sama” atau “anak bersikap sopan,” tapi tidak ada deskripsi perilaku yang jelas dan bisa diamati. Guru guru di TK IT Al Iffah juga masih bingung soal validitas dan reliabilitas instrumen. Banyak yang berpikir, selama semua indikator perkembangan anak sudah masuk, instrumennya menjadi bagus. Padahal, setiap indikator harus diuji terlebih dulu: valid atau tidak, stabil atau tidak. Dari sini terlihat, para guru masih perlu upgrade kapasitas dalam membuat sistem evaluasi yang benar. Jadi, waktu menyusun instrumen evaluasi di sekolah Islam seperti TK IT Al Iffah, tidak hanya fokus ke nilai akademik harus memasukan juga aspek spiritual, moral, sama sosial anak. Evaluasi yang tepat dan berkelanjutan itu kuncinya, supaya setiap anak benar-benar mendapat dukungan sesuai kebutuhannya. TK IT Al Iffah Bekasi sendiri merasa perlu memperbaiki sistem evaluasi yang ada, supaya lebih sesuai standar pendidikan modern yang menuntut hasil objektif dan personal. Saat ini, hasil penilaian masih sangat subjektif dan sulit dibandingkan antar-guru. Hal ini sejalan dengan temuan Mardapi¹⁰ yang menyampaikan apabila instrumen disusun tidak sistematis, maka penilaian bisa bias dan keadilan evaluasi juga berkurang.

METODE PENELITIAN

Penelitian ini menggunakan pendekatan kualitatif, karena fokusnya adalah memahami fenomena secara mendalam dan menyeluruh dalam konteks alami, yaitu bagaimana guru menggunakan dan mengembangkan instrumen evaluasi di TK IT Al Iffah. Pendekatan kualitatif deskriptif dipilih sebab penelitian ini tidak bermaksud menguji hipotesis atau menghasilkan angka statistik, melainkan untuk mendeskripsikan dan menganalisis rasionalitas instrumen yang digunakan oleh guru dalam praktik penilaian perkembangan anak usia dini. Peneliti

¹⁰ Djemari Mardapi, *Pengukuran Penilaian Dan Evaluasi Pendidikan* (Yigjakarta: Nuha Medika, 2018), <https://bintangpusnas.perpusnas.go.id/konten/BK59313/pengukuran-penilaian-dan-evaluasi-pendidikan>.

menjadi instrumen utama. Cara mengumpulkan datanya menggunakan triangulasi, lalu hasilnya dianalisis secara induktif. Intinya, penelitian ini lebih fokus nyari makna daripada sekadar menarik generalisasi.

Menurut Sugiyono, penelitian kualitatif adalah metode penelitian berbasis postpositivisme untuk meneliti kondisi objek alamiah, di peneliti jadi instrumen utama. Pengumpulan data menggunakan triangulasi, penelitian ini lebih fokus nyari makna daripada sekadar menarik generalisasi¹¹. Menurut Creswell¹², penelitian kualitatif bertujuan untuk mengeksplorasi dan memahami makna yang dianggap berasal dari masalah sosial atau kemanusiaan. Peneliti menjadi instrumen utama dalam pengumpulan data melalui interaksi langsung dengan subjek penelitian. Dalam konteks ini, peneliti berperan sebagai pengamat sekaligus penganalisis terhadap bentuk, isi, serta pelaksanaan instrumen evaluasi yang digunakan oleh guru di TK IT Al Iffah. Pendekatan deskriptif dipilih agar data yang diperoleh dapat menggambarkan kondisi nyata di lapangan secara naratif dan sistematis. Hasil penelitian ini diharapkan memberikan pemahaman yang lebih mendalam tentang sejauh mana rasionalitas instrumen yang digunakan sesuai dengan teori evaluasi pendidikan, baik dari segi validitas, reliabilitas, maupun kepraktisan.

Penelitian ini dilaksanakan di TK IT Al Iffah, sebuah lembaga pendidikan Islam terpadu yang berlokasi di kota Bekasi. Untuk memperoleh data yang akurat dan relevan, penelitian ini menggunakan tiga teknik pengumpulan data, yaitu: Observasi, Wawancara, Dokumentasi: Data dianalisis menggunakan model interaktif¹³ Miles & Huberman, yang terdiri atas tiga tahapan utama : Reduksi Data, Penyajian Data, Penarikan Kesimpulan dan Verifikasi: Kesimpulan ditarik berdasarkan temuan yang konsisten dari hasil triangulasi antara observasi, wawancara, dan dokumentasi.

Untuk menilai rasionalitas instrumen, peneliti menggunakan pendekatan analisis logis terhadap kesesuaian indikator dengan tujuan pembelajaran, kelengkapan rubrik penilaian, dan kesesuaian dengan karakteristik perkembangan anak usia dini.

¹¹ Sugiyono Sugiyono, *Metode Penelitian Kuantitatif, Kualitatif, Dan R&D* (Bandung: Alfabeta, 2019).

¹² John W. Creswell, *Research Design Qualitative and Quantitative Approaches*. (London: Sage Publication Internasional, 1994).

¹³ Roni K. Pratama, "Peran Penting Remaja demi Ketahanan Keluarga | Milenial," May 20, 2016, <https://www.gatra.com/news-426454-milenial-peran-penting-remaja-demi-ketahanan-keluarga.html>.

HASIL DAN PEMBAHASAN

Hasil penelitian menunjukkan bahwa TK IT Al Iffah menggunakan tiga jenis instrumen utama dalam mengevaluasi perkembangan peserta didik, yaitu lembar observasi perilaku, checklist perkembangan, dan portofolio hasil karya anak. Ketiga bentuk instrumen tersebut telah mencakup prinsip penilaian autentik (*authentic assessment*), yaitu menilai kemampuan anak dalam konteks nyata dan kegiatan sehari-hari. Namun, dari hasil analisis ditemukan bahwa penyusunan indikator belum sepenuhnya konsisten dengan tujuan pembelajaran mingguan. Sebagian item masih bersifat umum, misalnya “anak aktif bertanya” atau “anak peduli terhadap teman”, tanpa deskripsi perilaku yang dapat diamati secara spesifik. Ketiga instrumen ini mencerminkan upaya guru untuk menilai berbagai aspek perkembangan anak secara menyeluruh, baik dari dimensi kognitif, sosial-emosional, maupun spiritual.

1. Analisis Statistik dan Rasionalitas Instrumen

Tabel 1. Analisis rasional instrumen siswa TK IT AL IFFAH Bekasi

NAMA	Class	Gender	Age	Cognitive(0-10)	Language(0-10)	Motor(0-10)	Social(0-10)	Spiritual(0-10)	Behavior(1-4)
Arsyila Chalista	A	P	4	7.6	6.8	7.6	5.6	5.2	3
Al Farizki	B	L	4	6.5	6.5	7	7.8	4.3	1
Kirana Putri	B	P	5	7.6	5	5.9	5.8	5.8	3
Naufal Aldani	B	L	4	8.9	6.3	7.6	7.9	7.6	3
Alifya Falisha	A	P	4	7.2	6.4	6.4	8.4	6.7	2
Alfatunisa Assafiyah	A	P	5	8.9	5.8	8	6	4.7	3
M. Habibie Al ghifari	A	L	4	4.3	6.6	6.6	8.2	6.6	1
Aldi Syahputra	B	L	5	8.1	7.3	9	7.5	6.9	1
Nayaka Syakir	B	P	4	7.3	9.3	6.7	8	5.2	4
Alesha Zahra	B	P	5	6.9	7	7.2	9.3	6.6	3
Arevan Dirandra	A	L	5	7.3	7.1	8.3	6.7	6.5	2
A. Arkana Albiansyah	B	L	4	5	6.7	6.3	6.1	4.8	3
Mahira Nabila	B	P	4	7	4.3	7.7	5.9	6.9	3
Razka Rafif	A	L	5	7.6	6.8	8.8	6	7.2	2
Nabil Hafidz	A	L	5	8.8	6.9	5.9	6.9	7.9	3
Misyah Rahmadani	A	P	5	6.6	10	7.7	7.4	7.9	1
Meisya Ava	A	P	5	6.3	6.5	7.8	7.3	4.5	1
Rino Adihtama	B	L	4	6.6	7.2	8.3	8	5.1	3
Inas	A	P	4	8.2	6.8	6.3	7	7.1	3
Ellano	A	L	4	7.6	5.3	6.2	8.7	7.1	3
Gaffi	B	L	5	6.6	8.3	8	6.7	7.1	3
Nasyauqi	A	P	5	7.8	7.8	7.8	10	10	4
Haqi	A	L	4	7.3	7.8	7.8	7.8	7.2	3
Alfahri	A	L	4	8.3	5.6	7.8	6	8	2
Syakira Dera	A	P	4	6.4	8.6	6.8	5.7	7.7	3
Mirza Syafi	B	L	4	6.8	5	7.7	7.6	7.3	2
A. Fizah	A	P	5	6.8	7.6	7.8	6.7	6	3
M. Lutfi	B	L	5	5.6	9.6	6.8	7.9	7.5	1
Faro	B	L	5	7.5	5.5	9.4	7.6	5.3	1
Aqil M. Habibi	A	L	5	7.5	6.1	8	6.9	6.1	4

Berdasarkan hasil perhitungan dari data Excel (30 responden anak TK IT Al Iffah), diperoleh bahwa rata-rata skor tertinggi terdapat pada **aspek motorik (7,44)**, disusul aspek

sosial (7,25), kognitif (7,16), bahasa (6,88), dan spiritual (6,56). Hal ini menunjukkan bahwa perkembangan motorik anak lebih menonjol dibanding aspek spiritual.

Distribusi skor perilaku guru menunjukkan bahwa **mayoritas anak (50%) berada pada kategori “baik”**, sementara 23% anak masih memerlukan bimbingan dalam perilaku sosial dan disiplin. Ini memperlihatkan variasi kemampuan anak yang cukup lebar, yang seharusnya dapat menjadi dasar bagi guru dalam merancang pembelajaran diferensiatif.

Nilai **Cronbach’s Alpha sebesar 0,273** menandakan rendahnya konsistensi antar-aspek penilaian. Menurut Burhan Nurgiyantoro, hal ini bukan semata menunjukkan kelemahan instrumen, tetapi lebih menggambarkan bahwa lima domain perkembangan anak usia dini (kognitif, bahasa, motorik, sosial, spiritual) bersifat **multidimensional**, sehingga tidak diharapkan memiliki korelasi yang tinggi antar-aspek¹⁴. Meski demikian, hasil ini menegaskan pentingnya pembakuan rubrik penilaian agar perbedaan antar-guru dalam memberikan skor dapat diminimalkan. Sebagaimana dikemukakan oleh Mardapi, instrumen yang baik harus memenuhi tiga kriteria utama, yaitu **validitas, reliabilitas, dan praktikalitas**, agar hasil evaluasi dapat dipercaya dan digunakan sebagai dasar pengambilan keputusan pendidikan¹⁵.

2. Temuan Lapangan

Dari hasil wawancara, guru menyatakan bahwa mereka masih mengalami kesulitan dalam menyusun indikator penilaian yang spesifik dan terukur. Sebagian besar guru menyusun instrumen berdasarkan contoh dari lembaga lain atau pelatihan yang pernah diikuti. Hal ini menyebabkan beberapa instrumen tidak memiliki keterkaitan langsung dengan tujuan pembelajaran yang direncanakan. Selain itu, validitas isi (*content validity*) dari instrumen belum sepenuhnya diperhatikan. Misalnya, pada aspek spiritual, penilaian masih sebatas pada kehadiran anak dalam doa bersama, belum menyentuh dimensi internalisasi nilai-nilai agama. Guru juga menyadari bahwa belum ada pedoman observasi baku yang dapat dijadikan acuan bersama dalam melakukan penilaian.

Kendala lain adalah keterbatasan waktu guru dalam melakukan observasi secara mendalam terhadap seluruh anak setiap hari. Hal ini menyebabkan guru cenderung melakukan penilaian secara umum tanpa mencatat perilaku secara sistematis. Namun, upaya

¹⁴ Burhan Nurgiyantoro, *Penilaian Pembelajaran Bahasa Berbasis Kompetensi* (Bpfe Yogyakarta, 2016).

¹⁵ Mardapi, *Pengukuran Penilaian Dan Evaluasi Pendidikan*.

positif telah dilakukan sekolah melalui pelatihan internal tentang penyusunan rubrik penilaian yang lebih rinci dan terukur.

Temuan ini memperkuat hasil penelitian serupa disampaikan oleh Sukardi¹⁶ yang menyatakan bahwa guru sering kali mengandalkan pengalaman subjektif dalam melakukan evaluasi tanpa mempertimbangkan prinsip ilmiah pengukuran.

3. Pembahasan dan Implikasi

Secara umum, hasil penelitian menunjukkan bahwa instrumen evaluasi di TK IT Al Iffah telah mencakup seluruh aspek perkembangan anak, namun masih perlu peningkatan dalam hal rasionalitas dan konsistensi penyusunan indikator.

- Dari aspek validitas isi: beberapa indikator belum merepresentasikan kompetensi yang seharusnya diukur.
- Dari aspek reliabilitas: hasil penilaian antar-guru berpotensi berbeda karena belum adanya rubrik baku.
- Dari aspek praktikalitas: instrumen sudah mudah digunakan, tetapi belum efisien karena memerlukan waktu pengisian yang lama.

Guru perlu melakukan penyusunan ulang indikator berdasarkan prinsip evaluasi formatif yang dikemukakan oleh Arikunto¹⁷, yakni setiap instrumen harus mengukur hasil belajar yang sesuai dengan tujuan yang telah dirumuskan dalam rencana pembelajaran. Dengan demikian, rasionalitas instrumen di TK IT Al Iffah masih perlu ditingkatkan, terutama dalam aspek kejelasan indikator, validitas isi, dan keterpaduan dengan tujuan pembelajaran tematik. Langkah ke depan, sekolah dapat mengembangkan rubrik penilaian berbasis perilaku observabel agar hasil evaluasi lebih objektif dan konsisten antar-guru.

KESIMPULAN

Berdasarkan hasil penelitian mengenai analisis rasional instrumen dalam pengembangan sistem evaluasi di TK IT Al Iffah Bekasi, dapat disimpulkan bahwa secara umum lembaga telah memiliki sistem evaluasi yang berorientasi pada perkembangan menyeluruh anak, mencakup aspek kognitif, bahasa, motorik, sosial-emosional, dan spiritual. Guru juga telah menggunakan berbagai bentuk instrumen, seperti lembar observasi perilaku, checklist perkembangan, dan portofolio hasil karya anak, yang menunjukkan komitmen dalam menerapkan prinsip *authentic*

¹⁶ Sukardi, *Evaluasi Pendidikan*.

¹⁷ Arikunto, Supardi, and Suhardjono, *Penelitian Tindakan Kelas*.

assessment. Berdasarkan hasil perhitungan dari data Excel (30 responden anak TK IT Al Iffah), diperoleh bahwa rata-rata skor tertinggi terdapat pada **aspek motorik (7,44)**, Hal ini menunjukkan bahwa perkembangan motorik anak lebih menonjol dibanding aspek spiritual. Distribusi skor perilaku disiplin dan social anak menunjukkan bahwa **mayoritas anak (50%) berada pada kategori “baik”**, sementara 23% anak masih memerlukan bimbingan. Aspek validitas isi (*content validity*) juga masih perlu diperkuat. Sebagian item penilaian tidak sepenuhnya mewakili ranah kemampuan yang dinilai, khususnya dalam dimensi spiritual. Selain itu, nilai reliabilitas antar-guru (*inter-rater reliability*) masih rendah, sebagaimana terlihat dari hasil analisis statistik *Cronbach's Alpha* sebesar 0,273, yang menunjukkan bahwa setiap domain penilaian belum memiliki konsistensi internal yang tinggi. Kondisi ini dapat terjadi karena perbedaan persepsi antar-guru dalam memberikan skor dan belum adanya rubrik penilaian yang baku. Meskipun demikian, hasil penelitian juga menunjukkan adanya upaya positif dari pihak sekolah. Guru telah melakukan pelatihan internal untuk memperbaiki format penilaian dan mengembangkan rubrik yang lebih spesifik. Hal ini menunjukkan adanya kesadaran dan kemauan untuk memperbaiki sistem evaluasi agar lebih terstandar dan akuntabel. Rasionalitas instrumen evaluasi merupakan faktor kunci dalam menjamin kualitas hasil penilaian. Instrumen yang baik tidak hanya mencakup seluruh aspek perkembangan anak, tetapi juga harus memenuhi prinsip validitas, reliabilitas, objektivitas, dan praktikalitas. Dengan melakukan pembaruan sistem evaluasi berdasarkan teori ilmiah, sekolah seperti TK IT Al Iffah dapat memastikan bahwa setiap anak dinilai secara adil, objektif, dan sesuai dengan tahap perkembangannya. Evaluasi yang rasional akan menjadi fondasi penting dalam mewujudkan pendidikan Islam yang berkualitas, berkarakter, dan berorientasi pada pertumbuhan holistik anak.

DAFTAR PUSTAKA

- Andreella, Angela, Gaia Bertarelli, Federico Caldura, and Stefano Campostrini. “A Two-Step Model to Study the Inclusivity’s Distribution of Italian Early Childhood Education and Care Services.” Version 1. Preprint, arXiv, 2024. <https://doi.org/10.48550/ARXIV.2407.15610>.
- Arikunto, Suharsimi, Supardi, and Suhardjono. *Penelitian Tindakan Kelas: Edisi Revisi*. Jakarta: Bumi Aksara, 2015.
- Creswell, John W. *Research Design Qualitative and Quantitative Approaches*. London: Sage Publication Internasional, 1994.

- Gronlund, Norman Edward, and Robert L. Linn. *Measurement and Evaluation in Teaching*. Macmillan, 1990.
- Honghu, Yi, Liu Ting, and Lan Gongjin. "The Key Artificial Intelligence Technologies in Early Childhood Education: A Review." arXiv:2401.05403. Preprint, arXiv, December 20, 2023. <https://doi.org/10.48550/arXiv.2401.05403>.
- Khadka, Supriya, and Sanchari Das. "XR Design Framework for Early Childhood Education." arXiv:2601.18979. Preprint, arXiv, January 26, 2026. <https://doi.org/10.48550/arXiv.2601.18979>.
- Mardapi, Djemari. *Pengukuran Penilaian Dan Evaluasi Pendidikan*. Yigjakarta: Nuha Medika, 2018. <https://bintangpusnas.perpusnas.go.id/konten/BK59313/pengukuran-penilaian-dan-evaluasi-pendidikan>.
- Mulyasa. *Manajemen Paud*. Remaja Rosdakarya, 2014. <https://openlibrary.telkomuniversity.ac.id/pustaka/18540/manajemen-paud.html>.
- Nurgiyantoro, Burhan. *Penilaian Pembelajaran Bahasa Berbasis Kompetensi*. Bpfe Yogyakarta, 2016.
- Penchala, Sindhuja, Saketh Reddy Kontham, Prachi Bhattacharjee, Nima Mahmoodi, Daniel Fonseca, Sareh Karami, Mehdi Ghahremani, Andy D. Perkins, Shahram Rahimi, and Noorbakhsh Amiri Golilarz. "Learning in Focus: Detecting Behavioral and Collaborative Engagement Using Vision Transformers." Version 2. Preprint, arXiv, 2025. <https://doi.org/10.48550/ARXIV.2508.15782>.
- Pratama, Roni K. "Peran Penting Remaja demi Ketahanan Keluarga | Milenial." May 20, 2016. <https://www.gatra.com/news-426454-milenial-peran-penting-remaja-demi-ketahanan-keluarga.html>.
- Sugiyono, Sugiyono. *Metode Penelitian Kuantitatif, Kualitatif, Dan R&D*. Bandung: Alfabeta, 2019.
- Sukardi, Sukardi. *Evaluasi Pendidikan: Prinsip & Operasionalnya*. Jakarta: Bumi Aksara, 2017. <https://dpk.kepriprov.go.id/opac/detail/bv10f>.
- Yang, Yuanyuan, Yuan Shen, Tianchen Sun, and Yangbin Xie. *Validating the Effectiveness of a Large Language Model-Based Approach for Identifying Children's Development across Various Free Play Settings in Kindergarten*. Version 1. 2025. <https://doi.org/10.48550/ARXIV.2505.03369>.