

Perbandingan Skor Penilaian Konvensional dan Penilaian Berbasis Kahoot pada Materi Perubahan Wujud Benda Kelas IV MI

Puspadenty Arningtyas Anggraeny*

Pendidikan Guru Madrasah Ibtidaiyah, Sekolah Tinggi Agama Islam Grobogan, Indonesia

Abstrak

Penelitian ini membandingkan hasil penilaian siswa kelas IV MIU Jabalul Khoir 2 Purwodadi pada materi Perubahan Wujud Benda antara penilaian konvensional (tes tulis kertas) dan penilaian berbasis aplikasi Kahoot. Menggunakan pendekatan *mixed methods* dengan desain *sequential explanatory*, penelitian melibatkan 30 siswa kelas IV sebagai subjek penelitian. Data kuantitatif dikumpulkan melalui skor tes kedua instrumen, sedangkan data kualitatif diperoleh dari wawancara dan observasi. Integrasi data dilakukan melalui triangulasi pada tahap interpretasi, di mana temuan kualitatif digunakan untuk menjelaskan perbedaan skor antara kedua metode. Hasil analisis menunjukkan perbedaan skor rata-rata yang signifikan, dengan rata-rata skor Kahoot 78,4 dibandingkan penilaian konvensional 71,2. Uji *t* berpasangan menghasilkan nilai sig. 0,032 ($p < 0,05$). Temuan kualitatif mengungkap bahwa faktor motivasi, kecemasan tes, dan kondisi antarmuka berpengaruh pada perbedaan performa siswa. Penelitian ini merekomendasikan integrasi Kahoot sebagai instrumen penilaian formatif di madrasah ibtidaiyah, dengan catatan perlunya adaptasi terhadap keterbatasan infrastruktur teknologi.

Kata kunci: penilaian konvensional, penilaian Kahoot, perubahan wujud benda, mixed methods

Abstract

This study compares the assessment results of fourth-grade students of MIU Jabalul Khoir 2 Purwodadi on the material of Changes in the State of Objects between conventional assessment (paper-written test) and assessment based on the Kahoot application. Using a mixed methods approach with a sequential explanatory design, the study involved 30 fourth-grade students as research subjects. Quantitative data were collected through test scores of both instruments, while qualitative data were obtained from interviews and observations. Data integration was carried out through triangulation at the interpretation stage, where qualitative findings were used to explain the differences in scores between the two methods. The analysis results showed a significant difference in average scores, with an average Kahoot score of 78.4 compared to the conventional assessment of 71.2. The paired t-test produced a sig. 0.032 ($p < 0.05$). Qualitative findings revealed that motivational factors, test anxiety, and interface conditions influenced differences in student performance. This study recommends the integration of Kahoot as a formative assessment instrument in Islamic elementary schools, with a note on the need to adapt to the limitations of technological infrastructure.

Keywords: conventional assessment, Kahoot assessment, changes in the state of objects, mixed methods

Pendahuluan

Ada hal yang sering luput dari perhatian guru ketika menilai hasil belajar siswa, yaitu bahwa instrumen yang digunakan untuk mengukur kemampuan siswa bisa jadi turut memengaruhi kemampuan yang diukur itu sendiri. Seorang siswa yang paham betul dengan suatu materi belum tentu mampu mengungkapkan pemahamannya dalam bentuk soal esai atau pilihan ganda di atas selembar kertas dengan waktu terbatas dan suasana kelas yang tegang.

Asesmen pendidikan, dalam perspektif teori evaluasi pembelajaran, bukan sekadar alat pengukur hasil belajar. Menurut Bloom's Taxonomy of Educational Objectives yang telah menjadi

* Corresponding to the author: Puspadenty Arningtyas Anggraeny, Pendidikan Guru Madrasah Ibtidaiyah, Sekolah Tinggi Agama Islam Grobogan, Indonesia; e-mail: puspadenty.pd@gmail.com

landasan evaluasi pendidikan selama lebih dari tujuh dekade, penilaian seharusnya mampu menangkap berbagai level kognitif, dari pengetahuan faktual hingga kemampuan analisis dan sintesis (Bloom, 1956). Stiggins dalam teori *Student-Involved Classroom Assessment* menekankan bahwa instrumen asesmen yang baik tidak hanya mengukur apa yang siswa ketahui, tetapi juga mendorong motivasi intrinsik untuk belajar. Ketika format penilaian menciptakan kecemasan berlebih, yang terukur bukan lagi kemampuan kognitif murni, melainkan kemampuan kognitif yang tersaring oleh kondisi afektif (Stiggins, 2001).

Teori pengukuran klasik yang dikembangkan oleh Ibnu Hadjar menegaskan bahwa skor yang teramati merupakan fungsi dari skor sejati ditambah komponen kesalahan pengukuran. Dalam konteks perbandingan dua instrumen penilaian, perbedaan skor yang muncul bisa mencerminkan perbedaan skor sejati, tetapi bisa juga mencerminkan perbedaan komponen kesalahan yang berasal dari kondisi administrasi, format soal, dan kondisi psikologis siswa saat mengerjakan. Teori ini menjadi landasan untuk memahami mengapa dua instrumen yang dimaksudkan mengukur konstruk yang sama dapat menghasilkan skor yang berbeda (Prof. Dr. Ibnu Hadjar, 2022).

Teori motivasi belajar, khususnya *Self-Determination Theory* yang dikembangkan Deci dan Ryan, memberikan kerangka untuk memahami mengapa format penilaian berbasis gamifikasi seperti Kahoot dapat menghasilkan respons berbeda dari siswa. Teori ini membedakan motivasi ekstrinsik dari motivasi intrinsik dan menyatakan bahwa kondisi belajar yang mendukung otonomi dan kompetensi cenderung mengaktifkan motivasi intrinsik. Ketika siswa merasa penilaian adalah permainan bukan ujian yang menentukan nilai rapor, maka kondisi psikologis yang berbeda ini berpotensi melepaskan kapasitas kognitif yang dalam situasi penilaian konvensional terhambat oleh kecemasan (Edward L. Deci, 1985).

Dalam konteks IPA di sekolah dasar, Andayani dan Madani menegaskan bahwa penilaian yang baik harus valid, reliabel, dan adil. Artinya instrumen harus mengukur apa yang seharusnya diukur, konsisten dalam pengukurannya, dan tidak menguntungkan atau merugikan kelompok tertentu (Andayani & Madani, 2023). Nawali dkk. memandang bahwa penilaian seharusnya tidak hanya mengukur pengetahuan faktual tetapi juga pemahaman konseptual dan keterampilan proses sains (Jazilatun Nawali, Indah Aminatuz Zuhriyah, Samsul Susilawati, 2024). Materi Perubahan Wujud Benda di kelas IV, yang mencakup konsep mencair, membeku, menguap, mengembun, menyublim, dan mengkristal, semua berkaitan dengan transfer energi kalor yang bersifat abstrak bagi anak usia 9-10 tahun.

Sunaryati dkk. mencatat bahwa salah satu kelemahan penilaian konvensional untuk materi IPA di SD/MI adalah kecenderungannya untuk mengukur kemampuan hafalan lebih daripada pemahaman konsep. Siswa yang mahir menghafal nama-nama proses perubahan wujud bisa mendapat skor tinggi tanpa benar-benar memahami mengapa proses itu terjadi (Titin Sunaryati, Tazkia Aisha Laelly, Utari Febriyanti, Noviyanti, 2024). Di sisi lain, Rasyidin menemukan bahwa kecemasan tes secara konsisten berkorelasi negatif dengan performa siswa sekolah dasar pada penilaian tertulis, dengan intensitas yang lebih tinggi pada siswa introvert dan yang pernah mengalami kegagalan sebelumnya (Rasyidin, 2022). Dudin juga mencatat bahwa siswa dengan kemampuan membaca lebih rendah sering mengalami kesulitan memahami soal tertulis, sehingga skor mereka tidak mencerminkan pemahaman materi yang sebenarnya (Dudin, 2025).

Kemunculan platform digital dalam satu dekade terakhir menawarkan model penilaian alternatif. Aplikasi Kahoot, yang diluncurkan pada 2013 dan digunakan lebih dari 300 juta pengguna aktif di seluruh dunia, menggabungkan unsur gamifikasi ke dalam format kuis digital. Hartika

mencatat bahwa kombinasi hitungan mundur, efek suara, dan papan peringkat menciptakan pengalaman yang membuat siswa merasa lebih seperti bermain daripada diuji (Hartika, 2025). Shabiroh membandingkan Kahoot dengan Quizizz dan menyimpulkan bahwa Kahoot menghasilkan tingkat *engagement* yang lebih tinggi, meskipun kedua platform menghasilkan skor yang tidak berbeda signifikan (Shabiroh, 2023). Namun, Dewi mengingatkan bahwa fitur papan peringkat juga menciptakan tekanan kompetitif yang pada sebagian siswa justru meningkatkan kecemasan (Dewi, 2025).

Meskipun beberapa penelitian telah mengkaji efektivitas Kahoot, terdapat gap penelitian yang belum banyak dieksplorasi. Pertama, sebagian besar penelitian yang ada berfokus pada dampak Kahoot terhadap motivasi dan keterlibatan, bukan pada perbandingan skor kuantitatif yang langsung antara dua instrumen penilaian pada materi yang sama. Kedua, penelitian yang mengintegrasikan analisis kuantitatif dengan penjelasan kualitatif tentang mengapa skor berbeda masih sangat terbatas, terutama di konteks madrasah ibtidaiyah dengan keterbatasan infrastruktur teknologi. Ketiga, belum ada penelitian yang secara khusus menganalisis perbedaan manfaat Kahoot berdasarkan kelompok kemampuan siswa di konteks MI. Gap inilah yang menjadi dasar penelitian ini.

Penelitian ini bermula dari permasalahan di MIU Jabalul Khoir 2 Purwodadi, yang dalam dua tahun terakhir mulai mengintegrasikan teknologi ke dalam pembelajaran, namun belum menerapkan penilaian menggunakan aplikasi digital. Pembelajaran digitalisasi hanya berhenti pada ranah materi dan tidak dilanjutkan pada sistem penilaian. Hal ini menyebabkan kesenjangan penerapan penilaian, di mana siswa mendapatkan materi secara digitalisasi, akan tetapi proses penerapan penilaian kembali lagi ke desain konvensional. Sehingga perlu dilaksanakan proses penilaian dengan menggunakan aplikasi berbasis Kahoot supaya dapat diketahui perbedaan skor antara kedua metode penilaian tersebut. Azizah dalam pengembangan asesmen berbasis HOTS menggunakan Kahoot di SD menunjukkan bahwa instrumen tersebut sangat valid dan praktis, tetapi penelitiannya bersifat pengembangan, bukan komparatif terhadap penilaian konvensional (Azizah, 2023). Oleh karena itu, penelitian ini menjawab pertanyaan yang lebih spesifik, yaitu seberapa besar perbedaan skor antara kedua metode penilaian, dan apa yang menjelaskan perbedaan tersebut?

Metode

Penelitian ini menggunakan *Mixed methods research* dengan menggabungkan pendekatan kuantitatif dan kualitatif dalam satu kerangka studi. Penelitian ini mengidentifikasi pendekatan *mixed methods*, dengan desain *sequential explanatory*, di mana data kuantitatif dikumpulkan dan dianalisis terlebih dahulu, selanjutnya temuan kuantitatif dijelaskan melalui data kualitatif (Pane, 2021). Sugiyono menyatakan bahwa *mixed methods* relevan ketika peneliti ingin tidak hanya mengetahui ada tidaknya perbedaan (kuantitatif) tetapi juga mengapa perbedaan itu terjadi (kualitatif) (Sugiyono, 2019). Sejalan dengan Azhari dkk. penelitian ini menggunakan desain *sequential explanatory* untuk membandingkan efektivitas dua model pembelajaran dan menunjukkan bahwa temuan kualitatif kerap merevisi atau memperkaya interpretasi data kuantitatif yang tampaknya sudah jelas (Devi Syukri Azhari, Zihnil Afif, Martin Kustati, 2023).

Desain Penelitian ini menggunakan pendekatan *mixed methods* dengan desain *sequential explanatory*. Pada fase pertama, data kuantitatif dikumpulkan melalui administrasi dua instrumen asesmen kepada seluruh subjek penelitian. Pada fase kedua, temuan kuantitatif digunakan sebagai basis untuk mengembangkan pertanyaan wawancara dan panduan observasi yang lebih terfokus.

Subjek penelitian adalah seluruh siswa kelas IV MIU Jabalul Khoir 2 Purwodadi yang berjumlah 30 orang (17 laki-laki, 13 perempuan), dengan rentang usia 9-11 tahun. Karena populasi berjumlah kecil, digunakan teknik *total sampling* di mana semua anggota populasi dijadikan subjek penelitian. Instrumen Penilaian Konvensional yaitu Tes tulis terdiri dari 20 butir soal pilihan ganda (4 pilihan jawaban) dan 5 butir soal isian singkat yang mencakup seluruh sub-topik Perubahan Wujud Benda. Instrumen divalidasi oleh dua orang guru IPA berpengalaman dan satu dosen PGMI, kemudian diuji cobakan kepada 15 siswa kelas IV dari madrasah lain. Hasil uji reliabilitas menggunakan *Cronbach Alpha* menghasilkan nilai 0,79 (reliabilitas tinggi)

Instrumen Kahoot berupa Kuis Kahoot dirancang dengan 25 pertanyaan yang sama dengan tes konvensional, tetapi diadaptasi ke format pertanyaan tunggal dengan empat opsi jawaban dan durasi jawaban 20-30 detik per soal. Poin dihitung berdasarkan ketepatan dan kecepatan, kemudian dikonversi ke skala 100. Sebelum pelaksanaan, semua siswa diberi sesi latihan 5 menit menggunakan kuis Kahoot dengan contoh soal yang tidak berkaitan dengan materi.

Panduan wawancara semi-terstruktur untuk 8 siswa yang dipilih secara purposif (4 siswa dengan performa naik signifikan di Kahoot, 4 siswa dengan performa turun). Lembar observasi digunakan untuk mencatat perilaku siswa selama kedua sesi penilaian. Pengumpulan data berlangsung dalam dua sesi terpisah dengan jeda satu minggu untuk mengurangi efek belajar (*learning effect*). Sesi pertama menggunakan penilaian konvensional; sesi kedua menggunakan Kahoot. Urutan ini dipilih karena penilaian konvensional lebih familiar bagi siswa, sehingga sesi pertama tidak memberikan keuntungan adaptasi untuk siswa mana pun.

Teknik Analisis Data, Data kuantitatif dianalisis menggunakan uji *t* berpasangan (*paired sample t-test*) dengan perangkat lunak SPSS 24. Uji normalitas Shapiro-Wilk dilakukan sebelum uji *t* untuk memastikan asumsi parametrik terpenuhi. Data kualitatif dianalisis menggunakan analisis tematik dengan tiga tahap: kodifikasi awal, pengelompokan tema, dan interpretasi. Integrasi data *mixed methods* dilakukan pada tahap interpretasi hasil melalui proses triangulasi konvergen. Hasil analisis kualitatif digunakan secara sistematis untuk menjelaskan temuan kuantitatif yang menunjukkan perbedaan skor antara kedua metode penilaian. Secara konkret, setiap pola skor yang muncul dari analisis kuantitatif dikonfirmasi dan dijelaskan mekanismenya melalui tema-tema yang muncul dari wawancara dan observasi. Proses ini mengikuti model integrasi *merging* yang direkomendasikan Creswell, di mana data kuantitatif dan kualitatif dianalisis secara terpisah terlebih dahulu, kemudian dibandingkan dan diintegrasikan pada tahap pembahasan untuk memperoleh pemahaman yang lebih lengkap tentang fenomena yang diteliti (Creswell & Creswell, 2023). Triangulasi dilakukan dengan mencocokkan pola skor kuantitatif per kelompok siswa dengan narasi kualitatif yang diungkapkan dalam wawancara, sehingga penjelasan tentang perbedaan skor tidak bersandar semata pada angka statistik, melainkan diperkuat oleh perspektif siswa.

Hasil dan Diskusi

Hasil

1. Deskripsi Data Kuantitatif

Tabel berikut menyajikan statistik deskriptif hasil penilaian kedua instrumen:

Tabel 1. Statistik Deskriptif Hasil Penilaian

Statistik	Penilaian Konvensional	Aplikasi Kahoot
-----------	------------------------	-----------------

Rata-rata	71,2	78,4
Median	72,0	79,5
Std. Deviasi	12,3	10,8
Skor Minimum	40	52
Skor Maksimum	95	100
Siswa $\geq$ KKM (70)	18 (60%)	24 (80%)

Dari 30 siswa, 22 siswa (73,3%) mengalami peningkatan skor pada penggunaan Aplikasi Kahoot dibandingkan penilaian konvensional, 5 siswa (16,7%) mengalami penurunan skor, dan 3 siswa (10%) memiliki skor yang relatif sama (selisih < 3 poin).

2. Uji Normalitas dan Uji t Berpasangan

Uji normalitas Shapiro-Wilk menghasilkan nilai sig. 0,124 untuk data konvensional dan 0,087 untuk data Kahoot, keduanya di atas 0,05, sehingga asumsi normalitas terpenuhi. Uji t berpasangan menghasilkan:

Tabel 2. Hasil Uji t Berpasangan

Parameter	Nilai
t-hitung	2,28
Df	29
Sig. (2-tailed)	0,032
Mean difference	7,2
95% CI	[0,64; 13,76]

Karena $p = 0,032 < 0,05$, hipotesis nol (H_0 : tidak ada perbedaan signifikan) ditolak. Terdapat perbedaan yang signifikan secara statistik antara hasil penilaian konvensional dan penggunaan Aplikasi Kahoot pada taraf kepercayaan 95%.

3. Analisis Kelompok Berdasarkan Kemampuan Awal

Ketika data dianalisis berdasarkan kelompok kemampuan awal (tinggi, sedang, rendah berdasarkan nilai rapor semester sebelumnya), pola yang menarik muncul:

Tabel 3. Perbandingan Skor per Kelompok Kemampuan

Kelompok	n	Rata-rata Konvensional	Rata-rata Kahoot	Selisih
Tinggi	10	84,5	85,1	0,6
Sedang	12	70,3	78,7	8,4
Rendah	8	53,6	65,4	11,8

Pola ini menunjukkan bahwa penilaian menggunakan Aplikasi Kahoot lebih banyak memberikan keuntungan kepada siswa dengan kemampuan sedang dan rendah, sementara siswa berkemampuan tinggi tidak banyak terpengaruh oleh format penilaian. Perbedaan selisih yang mencolok antara kelompok rendah (11,8 poin) dan kelompok tinggi (0,6 poin) ini menjadi salah satu temuan paling substantif yang membutuhkan penjelasan kualitatif.

4. Temuan Kualitatif

Wawancara dengan delapan siswa dan observasi selama dua sesi penilaian menghasilkan tiga tema utama.

Tema 1: Kecemasan dan Relaksasi

Tujuh dari delapan siswa yang diwawancarai menyebutkan bahwa sesi Kahoot terasa tidak seperti ujian. Pada sesi konvensional, terpantau 4 siswa menggigit pensil, mengulang membaca soal berkali-kali, dan tampak tegang. Sedangkan pada sesi Kahoot, siswa tersebut lebih sering tersenyum, tidak menunjukkan kepanikan, dan beberapa menepuk meja saat berhasil menjawab dengan cepat.

Tema 2: Tekanan Waktu dan Dampaknya

Penilaian berbasis Kahoot juga memunculkan tekanan jenis lain, yaitu kecepatan menjawab. Beberapa siswa kehabisan waktu dalam proses pengerjaan. Hal ini mengonfirmasi bahwa format Kahoot tidak selalu adil bagi siswa dengan gaya belajar lebih reflektif.

Tema 3: Faktor Sosial dan Papan Peringkat

Empat dari delapan siswa menyebut papan peringkat sebagai sesuatu yang menegangkan sekaligus menambah semangat. Yang lebih menarik adalah bagaimana papan peringkat memengaruhi strategi siswa: dua siswa secara eksplisit mengaku lebih berani menebak saat tidak tahu jawabannya di Kahoot. Pada penilaian konvensional, keduanya cenderung mengosongkan jawaban yang tidak mereka yakini. Ini memengaruhi perbedaan skor secara langsung, terutama pada soal-soal yang secara konseptual tidak mereka kuasai.

Diskusi

1. Penggunaan Aplikasi Kahoot Menghasilkan Skor Lebih Tinggi

Hasil penelitian menunjukkan skor rata-rata Kahoot lebih tinggi 7,2 poin dari konvensional. Jika dikaji melalui kerangka pengukuran klasik, perbedaan ini tidak serta-merta berarti Kahoot lebih mudah atau lebih sulit, melainkan mencerminkan perbedaan komponen kesalahan pengukuran yang bersumber dari kondisi administrasi yang berbeda. Hal ini sejalan dengan temuan Suariantini dkk. yang menemukan bahwa instrumen asesmen numerasi menggunakan aplikasi Kahoot pada mata pelajaran matematika kelas IV di SD Gugus V Kecamatan Bebandem menunjukkan hasil validitas isi sebesar 1.0 yaitu berada pada kriteria validitas sangat tinggi, dan layak digunakan pada pembelajaran di SD (Ni Nyoman Gita Suariantini, Basilius Redan Werang, 2023).

Setidaknya ada tiga mekanisme yang bekerja berdasarkan integrasi temuan kuantitatif dan kualitatif: *Pertama*, pengurangan kecemasan tes. Sejalan dengan teori Stiggins tentang *Student-Involved Assessment*, suasana gamifikasi Kahoot terbukti menurunkan tingkat kecemasan tes, dan kecemasan yang lebih rendah memungkinkan siswa mengakses pengetahuan yang mereka miliki dengan lebih efektif. Ini bukan berarti skor Kahoot yang lebih tinggi adalah "inflasi", tapi bisa juga diargumentasikan bahwa skor konvensional dalam kondisi kecemasan adalah yang kurang valid sebagai cerminan kemampuan sebenarnya.

Kedua, efek tebak-tebakan yang lebih aktif. Temuan dari wawancara menunjukkan bahwa siswa lebih cenderung menebak di Kahoot daripada mengosongkan jawaban, karena Kahoot tidak memberikan penalti untuk jawaban salah. Hal ini secara mekanis meningkatkan peluang mendapat skor lebih tinggi, terutama untuk siswa yang memiliki pengetahuan parsial tentang suatu topik.

Ketiga, format pertanyaan yang lebih kontekstual. Dalam proses adaptasi soal untuk Kahoot, beberapa pertanyaan dirancang dengan skenario konkret yang pada penilaian konvensional disampaikan dengan bahasa lebih formal. Perubahan konteks penyajian ini berkontribusi pada peningkatan skor, terutama bagi siswa yang lebih responsif terhadap bahasa sehari-hari.

2. Perbedaan Skor dan Implikasi terhadap Validitas Instrumen

Perbedaan 7,2 poin signifikan secara statistik, tetapi apakah bermakna secara pedagogis? Dari perspektif psikometri, Subhaktiyasa mengingatkan bahwa dalam psikometri, dua instrumen dianggap ekuivalen jika menghasilkan skor yang berkorelasi tinggi dan tidak berbeda signifikan pada rata-ratanya (Subhaktiyasa, 2024). Dalam penelitian ini, meskipun korelasi antara kedua skor cukup tinggi ($r = 0,71$), perbedaan rata-rata yang signifikan menunjukkan bahwa Kahoot dan penilaian konvensional belum sepenuhnya ekuivalen. Data kualitatif memberikan penjelasan yang lebih kaya. Perbedaan skor tidak semata-mata mencerminkan perbedaan kemampuan kognitif, melainkan interaksi antara format penilaian dengan kondisi afektif siswa.

Dari perspektif penilaian formatif, yang lebih penting adalah apakah Kahoot mampu memberikan informasi yang akurat tentang area kelemahan siswa. Distribusi kesalahan jawaban di Kahoot menunjukkan pola yang konsisten dengan penilaian konvensional: soal-soal tentang sublimasi dan deposisi menjadi soal dengan persentase jawaban benar paling rendah di kedua instrumen. Konsistensi ini menunjukkan bahwa Kahoot tetap valid sebagai alat diagnostik meskipun rata-rata skornya lebih tinggi.

3. Pola Perbedaan Berdasarkan Kelompok Kemampuan: Integrasi Data

Temuan yang paling menarik secara pedagogis adalah pola berbeda pada tiga kelompok kemampuan. Kelompok rendah memperoleh keuntungan 11,8 poin di Kahoot, kelompok sedang 8,4 poin, sementara kelompok tinggi hanya 0,6 poin. Integrasi data kualitatif memberikan penjelasan mekanistik yang lebih lengkap. Siswa berkemampuan rendah pada penilaian konvensional cenderung membekukan diri ketika menghadapi soal tertulis, terpantau dari perilaku menggigit pensil atau mengulang membaca soal. Di Kahoot, kondisi permainan yang cair membuat mereka berani berpartisipasi, termasuk dengan menebak. Karena probabilitas tebakan benar pada soal empat pilihan adalah 25%, siswa yang sebelumnya mengosongkan seluruh soal yang tidak dikuasai akan mendapat rata-rata peningkatan skor yang cukup besar.

Siswa berkemampuan tinggi, sebaliknya, tidak banyak terpengaruh oleh format. Mereka sudah tahu jawabannya, sehingga format apapun tidak banyak mengubah akurasi mereka. Temuan ini konsisten dengan prediksi teori pengukuran klasik di mana pada siswa yang skor sejatinya sudah tinggi, komponen kesalahan pengukuran yang bersumber dari kondisi afektif relatif kecil dibandingkan siswa dengan skor sejati rendah.

4. Pertimbangan Khusus untuk Konteks Madrasah

Ada beberapa hal spesifik untuk konteks MI yang perlu mendapat perhatian tersendiri. *Pertama*, keterbatasan perangkat. Dari 30 siswa, 11 siswa tidak memiliki ponsel pintar sendiri dan menggunakan laptop milik sekolah. Empat siswa menyatakan kurang familiar dengan soal yang ada di layar, dan dua di antaranya termasuk kelompok yang skornya turun di Kahoot. Ini menunjukkan bahwa efek positif Kahoot bisa tereduksi oleh kendala akses teknologi.

Kedua, koneksi internet yang kurang stabil. Selama sesi Kahoot, terjadi dua kali gangguan koneksi yang menyebabkan sebagian siswa kehilangan waktu menjawab. Gangguan teknis ini

merupakan variabel yang tidak dapat dikontrol sepenuhnya dan membedakan kondisi penilaian digital dari kondisi penilaian konvensional yang lebih stabil.

Ketiga, kebaruan pengalaman. Semua siswa kelas IV MIU Jabalul Khoir 2 Purwodadi belum pernah menggunakan Kahoot sebelumnya untuk konteks penilaian formal. Antusiasme yang teramati bisa jadi sebagian besar didorong oleh *novelty effect*, bukan oleh sifat intrinsik Kahoot. Studi longitudinal dengan paparan Kahoot yang berulang diperlukan untuk memisahkan efek kebaruan dari efek gamifikasi yang sesungguhnya.

5. Relevansi dengan Kurikulum Merdeka

Kurikulum Merdeka mendorong penggunaan penilaian yang lebih variatif, autentik, dan formatif. Dalam kerangka ini, Kahoot relevan sebagai salah satu instrumen asesmen formatif, terutama untuk *quick check* pemahaman di akhir sesi pembelajaran (Arzaq, 2024). Kekuatan Kahoot justru terletak pada kemampuannya memberikan umpan balik instan kepada guru tentang soal mana yang paling banyak dijawab salah, yang dapat langsung digunakan untuk menyesuaikan strategi mengajar. Namun ia tidak dirancang sebagai satu-satunya instrumen penilaian sumatif. Temuan penelitian ini mendukung pendekatan komplementer: Kahoot untuk formatif, penilaian konvensional untuk sumatif.

Keterbatasan

Pertama, ukuran sampel yang terbatas. Penelitian ini hanya melibatkan 30 siswa dari satu kelas di satu madrasah. Meskipun total sampling digunakan, angka 30 tetap terlalu kecil untuk mengklaim generalisasi yang luas. Temuan penelitian ini paling tepat dibaca sebagai bukti awal (*preliminary evidence*) yang perlu dikonfirmasi dengan studi berskala lebih besar.

Kedua, urutan penilaian yang tidak dirotasi. Kedua sesi penilaian selalu dilaksanakan dalam urutan yang sama: konvensional lebih dulu, Kahoot setelahnya. Meskipun jeda satu minggu diberikan untuk meminimalkan efek belajar, urutan ini tetap menimbulkan potensi bias. Desain *counterbalanced* akan lebih tepat untuk mengendalikan ancaman validitas internal ini.

Ketiga, novelty effect yang belum bisa dipisahkan. Antusiasme yang teramati selama sesi Kahoot bisa jadi sebagian besar didorong oleh kebaruan pengalaman, bukan oleh sifat intrinsik Kahoot sebagai instrumen penilaian.

Keempat, konversi skor Kahoot yang tidak sempurna. Skor asli Kahoot dihitung berdasarkan kombinasi ketepatan dan kecepatan, kemudian dikonversi ke skala 100. Dua siswa yang sama-sama menjawab benar bisa mendapat skor berbeda hanya karena perbedaan kecepatan. Penyetaraan dua skor yang mengandung konstruk berbeda secara parsial adalah keterbatasan psikometri yang tidak bisa sepenuhnya dieliminasi.

Kelima, instrumen kualitatif yang terbatas pada delapan siswa. Ada kemungkinan bahwa siswa dengan perubahan skor moderat memiliki perspektif yang berbeda dan berpotensi memperkaya atau merevisi tema-tema yang ditemukan. Saturasi data kualitatif dalam penelitian ini belum sepenuhnya terjamin.

Keenam, faktor guru dan penyampaian instruksi. Kedua sesi difasilitasi oleh peneliti, bukan guru kelas reguler. Respons emosional siswa terhadap situasi penilaian bisa berbeda ketika yang memfasilitasi adalah wajah yang kurang familiar. Generalisasi temuan ke situasi di mana guru kelas sendiri yang menggunakan Kahoot memerlukan kehati-hatian.

Ketujuh, materi yang diuji hanya satu topik. Karakteristik materi Perubahan Wujud Benda yang berbasis konsep dan pilihan jawaban mungkin sangat cocok untuk format Kahoot. Materi IPA yang lebih prosedural atau abstrak mungkin tidak menghasilkan pola yang sama.

Kesimpulan

Penelitian ini menunjukkan bahwa terdapat perbedaan signifikan antara hasil penilaian siswa kelas IV MIU Jabalul Khoir 2 Purwodadi pada materi Perubahan Wujud Benda menggunakan penilaian konvensional (rata-rata 71,2) dibandingkan dengan menggunakan Kahoot (rata-rata 78,4), dengan nilai signifikansi $p = 0,032$.

Perbedaan ini tidak semata-mata mencerminkan bahwa Kahoot lebih mudah digunakan dalam penilaian. Yang lebih tepat adalah bahwa kedua instrumen mengukur konstruk yang sama tetapi dengan kondisi psikologis dan sosial yang berbeda. Perbedaan kondisi itulah yang menghasilkan perbedaan skor, sebagaimana dijelaskan oleh kerangka teori pengukuran klasik. Siswa berkemampuan sedang dan rendah mendapat manfaat paling besar dari format Kahoot (selisih 8,4 dan 11,8 poin), sementara siswa berkemampuan tinggi tidak banyak dipengaruhi (selisih 0,6 poin). Integrasi data kualitatif memperjelas mekanisme dibalik pola ini yaitu pengurangan kecemasan tes, perubahan strategi menjawab, dan dinamika sosial yang diciptakan oleh elemen gamifikasi semuanya berkontribusi pada perbedaan skor yang terpola berdasarkan kemampuan awal.

Kahoot merupakan instrumen asesmen yang berharga untuk diintegrasikan ke dalam praktik penilaian di MI sebagai alat penilaian formatif. Penilaian dengan menggunakan Kahoot tidak seharusnya menggantikan penilaian konvensional secara keseluruhan, tetapi keduanya dapat saling melengkapi. Kahoot dapat digunakan untuk memantau pemahaman secara cepat dan interaktif, sedangkan penilaian konvensional tetap diperlukan untuk asesmen sumatif yang memerlukan dokumentasi formal dan kondisi yang terstandarisasi.

Referensi

- Andayani, T., & Madani, F. (2023). Peran Penilaian Pembelajaran Dalam Meningkatkan Prestasi Siswa di Pendidikan Dasar. *Jurnal Educatio*, 9(2), 924–930. <https://doi.org/10.31949/educatio.v9i2.4402>
- Arzaq, M. Y. (2024). *Inovasi Pembelajaran SD/MI Berbasis Kurikulum Merdeka di Era Digital* (1st ed.). PT NAsya Expanding management.
- Azizah, E. (2023). *Pengembangan Asesmen Berbasis Hots Menggunakan Aplikasi Kahoot Pada Mata Pelajaran Ipas Kelas Iv Sekolah Dasar*. repository.unja.ac.id. <https://repository.unja.ac.id/45950/>
- Bloom, B. S. (1956). *Taxonomy of Educational Objectives: The Classification of Educational Goals* (vol. 1). D. McKay.
- Creswell, J. W., & Creswell, J. D. (2023). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. SAGE. <https://books.google.co.id/books?id=b9tFzwEACAAJ>
- Devi Syukri Azhari, Zihnil Afif, Martin Kustati, N. S. (2023). Penelitian Mixed Method Research Untuk Disertasi. *INNOVATIVE: Journal Of Social Science Research*, 3(2), 8010–8025. <https://j-innovative.org/index.php/Innovative/article/view/1339/979>
- Dewi, F. (2025). *Pengembangan Media Pembelajaran Kahoot Pada Pembelajaran Bahasa Indonesia Kelas V Sd/Mi*. repository.radenintan.ac.id. <https://repository.radenintan.ac.id/37976/>
- Dudin, K. I. (2025). *Analisis Kesulitan Siswa Dalam Menyelesaikan Soal Literasi Sains Di Mi Negeri 3 Purbalingga*. <https://repository.uinsaizu.ac.id/id/eprint/34693>

- Edward L. Deci, R. M. R. (1985). *Intrinsic Motivation and Self-Determination in Human Behavior* (1st ed.). Springer New York, NY. <https://doi.org/https://doi.org/10.1007/978-1-4899-2271-7>
- Hartika, R. (2025). *Aplikasi Kahoot Sebagai Media Pembelajaran Bahasa Arab Untuk Meningkatkan Motivasi Belajar Peserta Didik Di SD Islam Terpadu Insan Gemilang Sigi*. repository.uindatokarama.ac.id. <https://repository.uindatokarama.ac.id/id/eprint/4657/>
- Jazilatun Nawali, Indah Aminatuz Zuhriyah, Samsul Susilawati, A. Z. N. Y. (2024). Implementasi Penilaian Autentik Di Sdi Surya Buana Malang Untuk Meningkatkan Kualitas Pembelajaran. *Pendas : Jurnal Ilmiah Pendidikan Dasar*, 09(04), 2548–6950. <https://doi.org/https://doi.org/10.23969/jp.v9i04.20886>
- Ni Nyoman Gita Suariantini, Basilius Redan Werang, I. G. A. (2023). Instrumen Asesmen Numerasi Online Menggunakan Aplikasi Kahoot Pada Mata Pelajaran Matematika Kelas IV Sekolah Dasar. *Innovative: Journal Of Social Science Research*, 3(2), 5712–5824. <http://j-innovative.org/index.php/Innovative/article/view/930>
- Pane, I. (2021). *Desain penelitian mixed method*. Yayasan Penerbit Muhammad Zaini.
- Prof. Dr. Ibnu Hadjar, M. E. (2022). *Teori Klasik Dan Modern dalam Pengembangan Tes*. CV. Eureka Media Aksara.
- Rasyidin, M. U. (2022). Penerapan Teknik Relaksasi Untuk Menangani Kecemasan Menghadapi Ujian Pada Siswa. *Jurnal BK Unesa*, 12(2), 865–877. <https://ejournal.unesa.ac.id/index.php/jurnal-bk-unesa/article/view/45875>
- Shabiroh, S. A. U. (2023). Komparasi Penggunaan Media Pembelajaran Quizizz Dan Media Pembelajaran Kahoot Terhadap Hasil Belajar Sejarah Siswa Kelas Xii Ipa Sma Negeri 1 Balongpanggang. *Avatara: Jurnal Pendidikan Sejarah*, 13(2), 0–7. <https://ejournal.unesa.ac.id/index.php/avatara/article/view/51348>
- Stiggins, R. J. (2001). *Student-Involved Classroom Assessment* (3rd ed.). Merrill Prentice Hall.
- Subhaktiyasa, P. G. (2024). Evaluasi Validitas dan Reliabilitas Instrumen Penelitian Kuantitatif: Sebuah Studi Pustaka. *Journal of Education Research*, 5(4), 5599–5609. https://www.researchgate.net/profile/Putu-Gede-Subhaktiyasa/publication/386012499_Evaluasi_Validitas_dan_Reliabilitas_Instrumen_Penelitian_Kuantitatif_Sebuah_Studi_Pustaka/links/673f3f0bb2d7de5adc4a0ed2/Evaluasi-Validitas-dan-Reliabilitas-Instrumen-Penelitian-Kuantitatif-Sebuah-Studi-Pustaka.pdf
- Sugiyono, P. D. (2019). *Metode penelitian pendidikan : (pendekatan kuantitatif, kualitatif dan R & D)*. Alfabeta Bandung.
- Titin Sunaryati, Tazkia Aisha Laelly, Utari Febriyanti, Noviyanti, F. A. (2024). Analisis Komprehensif Terhadap Jenis-Jenis Evaluasi Pembelajaran Sekolah Dasar: Tinjauan Literatur. *INNOVATIVE: Journal Of Social Science Research*, 4(04), 32435–32444. <https://doi.org/https://doi.org/10.31004/innovative.v4i4.13463>