

PENGEMBANGAN ASESMEN NASIONAL BERBASIS KELOKALAN

Badrun Kartowagiran

Pascasarjana Universitas Negeri Yogyakarta
kartowagiran@uny.ac.id

ABSTRAK

Pengembangan asesmen nasional berbasis kelokalan. Asesmen nasional itu adalah AKM yang dilaksanakan di kelas 5 atau di kelas 2 untuk SLTP dan SLTA. Hal ini dapat dipahami karena ujian nasional di akhir jenjang itu sudah berakhir pada tahun 2019. Pertanyaan lainnya adalah sifat kelokalan untuk asesmen nasional. Dalam pikiran penulis yang dimaksud kelokalan pada asesmen nasional adalah lokal Indonesia, bukan lokal daerah karena soal asesmen ini akan digunakan di seluruh wilayah Indonesia. Apabila butir soal asesmen nasional ini mengandung local daerah maka dapat dikatakan butir soal itu bias budaya.

Ujian nasional berbasis kelokalan dalam tulisan ini adalah asesmen kompetensi yang berbasis pada Bhineka Tunggal Ika. Soal yang digunakan untuk penilaian dapat berbentuk pilihan ganda, pilihan ganda kompleks, isian, dan uraian. Soal harus berkualitas tinggi dilihat dari aspek materi, konstruksi, dan bahasa. Soal juga harus mendidik peserta tes untuk berfikir kritis, memecahkan masalah, dan kreatif.

Kata kunci: *Asesmen nasional, kelokalan*

PENDAHULUAN

Pengembangan asesmen nasional berbasis kelokalan. Terhadap judul ini muncul pertanyaan apakah yang dimaksud dengan asesmen nasional itu adalah asesmen kompetensi minimum (AKM)? Untuk sementara penulis anggap bahwa yang dimaksud asesmen nasional itu adalah AKM yang dilaksanakan di kelas 5 atau di kelas 2 untuk SLTP dan SLTA. Hal ini dapat dipahami karena ujian nasional di akhir jenjang itu sudah berakhir pada tahun 2019. Pertanyaan lainnya adalah sifat kelokalan untuk asesmen nasional. Dalam pikiran penulis yang dimaksud kelokalan pada asesmen nasional adalah lokal Indonesia, bukan lokal daerah karena soal asesmen ini akan digunakan di seluruh wilayah Indonesia. Apabila butir soal asesmen nasional ini mengandung local daerah maka dapat dikatakan butir soal itu bias budaya.

Pertanyaan yang muncul lainnya adalah terkait dengan istilah asesmen; apakah asesmen di sini sama dengan penilaian? Hal ini perlu dipertanyakan karena menurut penulis, ada sedikit perbedaan antara penilaian dan asesmen. Penilaian itu ada yang sumatif (misal ujian akhir semester dan AKM) dan ada yang formatif. Lain halnya dengan istilah asesmen yang selalu formatif. Untuk memudahkan pemahaman pembaca, dalam kesempatan ini penulis setuju menyamakan asesmen dengan penilaian.

HASIL DAN PEMBAHASAN

Pada pembahasan ini akan dibagi menjadi dua, yakni: (1) pengertian penilaian, (2) prinsip-prinsip penilaian, dan (3) cara mengembangkan asesmen nasional berbasis kelokalan.

1. Pengukuran, penilaian, dan evaluasi

Penilaian merupakan sebuah istilah yang lazim ditemukan dan didengar dalam dunia pendidikan. Rangkaian kegiatan penilaian dilaksanakan untuk memperoleh dan menafsirkan data tentang proses dan hasil belajar peserta didik yang dilakukan secara sistematis dan berkesinambungan. Data yang berhasil dikumpulkan menggunakan alat atau instrumen yang telah disusun sesuai dengan karakteristik materi yang akan dikumpulkan. Secara umum, masyarakat awam yang menerima laporan hasil belajar peserta didik, misalnya orang tua atau wali murid, akan lebih mengenal istilah penilaian dibandingkan dengan pengukuran dan evaluasi. Istilah pengukuran, evaluasi, dan penilaian sering sekali tumpang tindih atau justru disamakan, padahal ketiganya adalah tiga hal yang berbeda. Oleh karena itu, dalam makalah ini akan dijelaskan perbedaan antara pengukuran, penilaian, dan evaluasi.

Terkait dengan pengukuran, Gronlund (1985) menyatakan "*measurement as process of obtaining of numerical description of the degree to which an individual pocesses a particular characteristic*". Selanjutnya Keeves & Masters (1999) juga mengatakan bahwa pengukuran adalah pemberian angka (kuantitas numerik) pada objek-objek atau kejadian-kejadian menurut aturan. Pengukuran terdiri dari aturan-aturan untuk memberikan angka/bilangan kepada objek dengan cara yang sedemikian rupa sehingga dapat mempresentasikan secara kuantitatif sifat-sifat objek tersebut (Nunnally: 1978). Definisi pengukuran yang dijelaskan para ahli di atas menegaskan bahwa dalam pemberian angka pada subjek, objek atau kejadian tidak asal memberi angka namun harus menggunakan aturan-aturan. Artinya, orang yang akan memberi angka pada subjek, objek, ataupun kejadian harus memperhatikan kaidah-kaidah tertentu agar angka yang diberikan sesuai dengan keadaan yang sebenarnya. Semakin jauh seseorang meninggalkan aturan-aturan pengukuran maka semakin besar kesalahan yang terjadi. Dalam hal ini, dapat disimpulkan bahwa pengukuran adalah kegiatan secara sistematis untuk mengkuantifikasikan (menganangkakan) suatu subyek atau obyek atau *trait* (sifat).

Pengukuran itu sendiri dapat dilakukan melalui tes dan nontes. Menurut para ahli, *testing* adalah suatu proses pengangkaan atau kuantifikasi potensi kognitif dan psikomotor dengan menggunakan alat yang dirancang secara khusus. Ujian bagi peserta didik merupakan proses kuantifikasi prestasi belajar peserta didik dalam bidang tertentu melalui tes, sedangkan mengukur panjang meja atau tinggi badan seseorang merupakan kuantifikasi suatu objek melalui nontes. Secara sederhana, Salkind (2013) menjelaskan tes adalah kegiatan sistematis yang digunakan untuk mengetahui tingkat penguasaan seseorang pada bidang atau keterampilan tertentu. Sementara itu Allen & Yen (1979) menyebut tes sebagai "*device for obtaining a sampel of an individual's behavior*". Tes merupakan instrumen atau prosedur sistematis untuk mengukur sampel perilaku seseorang. Dalam kasus tertentu ujian sama dengan tes, yakni suatu proses yang sistematis dalam mengkuantifikasikan potensi kognitif dan psikomotor dengan menggunakan alat yang dirancang secara khusus.

Sementara itu yang dimaksud dengan penilaian adalah menafsirkan atau membandingkan skor (bisa juga hasil Pengukuran) yang implikasinya dikenakan pada orang perorang, misal kamu lulus atau kamu belum lulus. Atau boleh juga dijelaskan bahwa penilaian adalah kegiatan sistematis untuk mengumpulkan, menganalisis, dan menyajikan informasi secara akurat dan bermanfaat untuk menafsirkan keberhasilan belajar peserta didik. Stiggins (2012) menjelaskan bahwa penilaian adalah proses pengumpulan informasi tentang

belajar peserta didik untuk perbaikan pembelajaran. Wright dan Stones (1992) menuliskan "*assessment provides an accounting of how much student learn in school and what resources are expended on achieving those learning outcome*". Penilaian dapat menjelaskan seberapa jauh peserta didik belajar di sekolah dan sumber apa saja yang diperlukan untuk mencapai hasil pembelajaran tersebut. Sementara itu, Djemari Mardapi (2008) menjelaskan bahwa penilaian meliputi pengumpulan bukti-bukti tentang pencapaian belajar peserta didik. Bukti ini tidak selalu diperoleh melalui tes, tetapi juga bisa dikumpulkan melalui pengamatan atau laporan diri. Dengan demikian yang dimaksud dengan penilaian dalam buku ini adalah suatu proses yang sistematis dan mencakup kegiatan mengumpulkan, menganalisis, serta menginterpretasikan informasi untuk menentukan capaian peserta didik/subyek yang dinilai terhadap tujuan pembelajaran yang telah ditetapkan.

Selanjutnya hasil penilaian ini digunakan untuk melakukan evaluasi, utamanya evaluasi pembelajaran. Apabila 90 % siswa dalam suatu kelompok pembelajaran lulus dan kriteria keberhasilan program 76% maka dapat dikatakan bahwa program pembelajaran itu berhasil. Secara umum dapat dikatakan bahwa evaluasi merupakan suatu proses menyediakan informasi yang dirancang untuk membantu dalam pengambilan keputusan tentang objek yang dievaluasi, sedangkan objek evaluasi tersebut berupa perencanaan, program, kebijakan, organisasi, produk, maupun individual (Owen: 1993). Worthen dan Sanders (1973) berpendapat "*Evaluation is the determination of the worth of a thing. It includes obtaining information for use in judging the worth of a program, product, procedure, or objective, or the potential utility of alternative approaches designed to attain specified objective*". Dalam pengertian ini evaluasi merupakan penentuan nilai sesuatu. Evaluasi meliputi pengumpulan informasi yang digunakan untuk mempertimbangkan nilai suatu program, produk, prosedur, tujuan atau manfaat potensial dari pendekatan alternatif yang di desain untuk mencapai tujuan yang telah ditetapkan. Pendapat tersebut mengimplikasikan adanya kriteria tertentu yang digunakan untuk menentukan nilai serta adanya sesuatu yang dinilai. Kriteria yang dimaksudkan adalah keberhasilan pelaksanaan program, dan hal yang dinilai adalah dapat berupa dampak atau hasil yang dicapai atau prosesnya sendiri.

Fitzpatrick, Sanders, dan Worthen (2004) menjelaskan evaluasi adalah proses pengumpulan informasi melalui inquiry atau teknik lain yang relevan, membandingkan dengan kriteria untuk menentukan nilai, kualitas, kegunaan, efektivitas, atau signifikansi. Selanjutnya Stufflebeam & Shinkfield (2007) mendefinisikan *evaluation is a process for giving attestation on such matters as reliability, effectiveness, cost-effectiveness, efficiency, safety, ease of use, and probity*. Dari definisi tersebut, diketahui bahwa evaluasi merupakan suatu proses untuk memberikan pengesahan pada hal-hal seperti keandalan, efektivitas, efektivitas biaya, efisiensi, keamanan, kemudahan penggunaan, dan kejujuran.

Berdasarkan beberapa pendapat tersebut dapat disimpulkan bahwa evaluasi di bidang pendidikan adalah suatu proses pengumpulan, pengolahan, dan analisis data yang dilakukan secara sistematis untuk melihat sejauh mana kebijakan/program/proyek tersebut telah mencapai tujuan berdasarkan kriteria-kriteria yang telah ditentukan sebelumnya. Hasil evaluasi tersebut dapat dijadikan dasar untuk menentukan nilai manfaat atau keberhasilan suatu program pendidikan, yang diwujudkan dalam bentuk rekomendasi dari evaluator

kepada pengambil keputusan dalam pengambilan keputusan tentang penyempurnaan program, peningkatan program ke arah yang lebih baik, perluasan implementasi program, maupun penghentian program. Evaluasi juga dapat memberikan informasi optimal yang dapat digunakan untuk meningkatkan kualitas pembelajaran atau pendidikan serta pengambilan kebijakan.

Uraian di atas menegaskan bahwa terdapat perbedaan antara evaluasi, penilaian, pengukuran, dan tes. Pengukuran adalah kegiatan secara sistematis untuk mengkuantifikasikan (mengangkakan) suatu subyek atau obyek atau *trait* (sifat). Pengukuran dapat dilakukan melalui tes dan non tes, yang dalam hal ini tes sebagai kata kerja. Tes sebagai kata kerja merupakan kegiatan secara sistematis untuk mengkuantifikasikan suatu potensi, sedangkan tes sebagai kata benda adalah seperangkat pertanyaan yang jawabannya ada unsur benar dan salah. Penilaian merupakan penafsiran terhadap informasi/skor hasil pengukuran yang implikainya dikenakan pada orang perorang. Sementara itu, evaluasi penafsiran terhadap informasi/skor hasil penilaian yang implikasinya dikenakan pada sekelompok orang atau program. Dengan demikian dapat disimpulkan bahwa tes sama dengan ujian, sedangkan tes bagian dari pengukuran, pengukuran bagian dari penilaian, dan penilaian bagian dari evaluasi.

2. Prinsip-prinsip penilaian

Untuk menjamin mutu pelaksanaan penilaian di jenjang pendidikan dasar dan menengah, pendidik dan pemangku kepentingan harus memperhatikan prinsip-prinsip penilaian seperti yang tertera dalam Standar Penilaian atau Kepmen Nomor 23 Tahun 2016 berikut.

- a. Sahih, berarti penilaian didasarkan pada data yang mencerminkan kemampuan yang diukur.
- b. Objektif, berarti penilaian didasarkan pada prosedur dan kriteria yang jelas, tidak dipengaruhi subjektivitas penilai.
- c. Adil, berarti penilaian tidak menguntungkan atau merugikan peserta didik karena berkebutuhan khusus serta perbedaan latar belakang agama, suku, budaya, adat istiadat, status sosial ekonomi, dan gender.
- d. Terpadu, berarti penilaian oleh pendidik merupakan salah satu komponen yang tak terpisahkan dari kegiatan pembelajaran.
- e. Terbuka, berarti prosedur penilaian, kriteria penilaian, dan dasar pengambilan keputusan dapat diketahui oleh pihak yang berkepentingan.
- f. Menyeluruh dan berkesinambungan, berarti penilaian oleh pendidik mencakup semua aspek kompetensi dengan menggunakan berbagai teknik penilaian yang sesuai, untuk memantau perkembangan kemampuan peserta didik.
- g. Sistematis, berarti penilaian dilakukan secara berencana dan bertahap dengan mengikuti langkah-langkah baku.
- h. Beracuan kriteria, berarti penilaian didasarkan pada ukuran pencapaian kompetensi yang ditetapkan.
- i. Akuntabel, berarti penilaian dapat dipertanggungjawabkan, baik dari segi teknik, prosedur, maupun hasilnya.

Agar penilaian memenuhi prinsip-prinsip penilaian, maka instrumen harus dikembangkan dengan baik. Terkait dengan hal ini, dalam Ayat 5, Permendiknas Nomor 20 Tahun 2007 dijelaskan bahwa instrumen penilaian hasil belajar yang digunakan pendidik memenuhi persyaratan (a) substansi, adalah merepresentasikan kompetensi yang dinilai, (b) konstruksi, adalah memenuhi persyaratan teknis sesuai dengan bentuk instrumen yang digunakan, dan (c) bahasa, adalah menggunakan bahasa yang baik dan benar serta komunikatif sesuai dengan taraf perkembangan peserta didik.

3. Pengembangan Asesmen Nasional Berbasis Kelokalan

Dalam pembahasan ini akan disampaikan jenis penilaian, pengertian kelokalan, dan cara mengembangkan butir soal yang berbasis kearifan lokal.

a. Jenis penelitian

Secara garis besar, penilaian dapat dikelompokkan menjadi dua, yakni penilaian formatif dan penilain sumatif. Pemisahan kata formatif dan sumatif dilakukan pertama kali oleh Scriven (1967). Lebih jauh Scriven menjelaskan bahwa evaluasi sumatif terjadi ketika para pendidik mengevaluasi sebuah program pembelajaran secara tuntas. Sebaliknya, evaluasi formatif terjadi ketika para pendidik mengevaluasi sebagian program pembelajaran. Evaluasi sumatif digunakan untuk menginformasikan keputusan *go/no-go* tentang suatu program, sedangkan evaluasi formatif digunakan untuk membuat keputusan perbaikan terkait suatu program. Sementara itu, Dunn dan Mulvenon (2009: 4) menyatakan bahwa evaluasi formatif pada pembelajaran adalah evaluasi berdasarkan bukti dari data penilaian (asesmen) agar guru, siswa, dan pemangku kepentingan pendidikan memperoleh informasi tentang proses belajar mengajar.

Saat ini istilah formatif dan sumatif itu juga digunakan oleh guru dalam melakukan penilaian. Menurut Popham (2017) penilaian formatif adalah proses terencana di mana bukti yang diperoleh melalui penilaian mengenai status siswa digunakan oleh guru untuk menyesuaikan prosedur pembelajaran mereka yang sedang berlangsung atau oleh siswa untuk menyesuaikan taktik belajar mereka saat ini. Penilaian formatif bukanlah tes melainkan proses yang disusun dengan cermat di mana hasil penilaian digunakan oleh guru atau siswa untuk meningkatkan apa yang mereka lakukan. Proses pengumpulan bukti berbasis penilaian tersebut, kemudian menggunakan bukti ini untuk membuat penyesuaian yang diperlukan, adalah jantung dari proses penilaian formatif. Ini adalah strategi untuk menggunakan bukti yang diperoleh dari penilaian untuk meningkatkan pengajaran guru dan untuk meningkatkan pembelajaran siswa.

Perencanaan yang matang perlu dilakukan sejak awal dalam proses penilaian formatif; suatu proses yang jelas bukan merupakan upaya yang berpatokan pada waktu. Guru dapat saja, membuat perubahan di tempat setiap saat. Misalnya, di tengah-tengah guru menjelaskan tentang konsep tertentu, jika guru melihat bahwa sebagian besar siswa terlihat benar-benar bingung, guru mungkin akan membuat keputusan instan untuk memulai penjelasan lagi, dan mencoba untuk meningkatkan kejernihannya dalam melakukan penjelasan. Perubahan strategi pembelajaran secara spontan semacam ini baik-baik saja atau tidak masalah bahkan harus didorong agar siswa tidak mengalami kebingungan.

Di bagian lain, Popham (2017) menjelaskan bahwa *summative assessment* terjadi ketika pendidik mengumpulkan bukti berbasis tes untuk menginformasikan keputusan tentang kegiatan pembelajaran yang sudah selesai seperti ketika tes akuntabilitas di seluruh provinsi atau negara yang diberikan setiap tahun untuk menentukan efektivitas pembelajaran sekolah selama satu tahun. Sebagai contoh penilaian sumatif lainnya, adalah ujian akhir yang diberikan guru kepada murid-muridnya. Tujuan ujian itu adalah untuk membantu guru memberikan nilai yang sesuai untuk murid-muridnya berdasarkan pada seberapa banyak yang telah dipelajari siswa selama satu semester. atau tahun sekolah. Keputusan yang didasarkan pada hasil kinerja siswa pada ujian akhir bukanlah keputusan peningkatan pengajaran, melainkan keputusan tentang nilai apa yang harus diberikan guru kepada murid-muridnya.

Berbeda dengan penilaian sumatif, penilaian formatif berfokus pada peningkatan cara guru mengajar sesuatu dan cara siswa belajar sesuatu. Penilaian formatif memiliki misi pengukuran "membuat lebih baik", sedangkan penilaian sumatif mencoba menjawab pertanyaan "kualitas pembelajaran". Menurut McMillan (2020) penilaian formatif itu seperti halnya *the chef tastes the soup*, sedangkan penilaian sumatif seperti halnya *the guests taste the soup*. Ringkasnya, penilaian sumatif digunakan untuk mengetahui kualitas pembelajaran atau kualitas lulusan dan salah satu asesmen yang termasuk penilaian sumatif adalah Asesmen Nasional. Pasal 8, Permendikbud Nomor 23 Tahun 2016 menjelaskan bahwa Ujian Nasional digunakan untuk: (a) pemetaan mutu program dan/atau satuan pendidikan; (b) pertimbangan seleksi masuk ke jenjang pendidikan berikutnya; dan (c) pembinaan dan pemberian bantuan kepada satuan pendidikan dalam upayanya untuk meningkatkan mutu pendidikan.

4. Ujian nasional yang berbasis kelokalan

Secara umum dapat dijelaskan bahwa kearifan lokal (*local wisdom*) itu merupakan pandangan hidup, ilmu pengetahuan, dan berbagai strategi kehidupan yang berwujud aktivitas yang dilakukan oleh masyarakat setempat untuk menjawab berbagai masalah dalam pemenuhan kebutuhan mereka. Disamping itu kearifan lokal dapat pula dimaknai sebagai sebuah sistem dalam tatanan kehidupan sosial, politik, budaya, ekonomi, dan lingkungan yang hidup di dalam masyarakat lokal. Karakter khas yang *inherent* dalam kearifan lokal sifatnya dinamis, kontinu, dan diikat dalam komunitasnya.

Di beberapa daerah yang ada di wilayah Indonesia juga memiliki kearifan local yang bersifat slogan positif, misalnya: (1) nilai yang terkandung dalam semboyan "*heuras peureupna, pageuh keupeulna tur lega awurna*" telah mampu memotivasi orang sunda untuk tampil sebagai pekerja keras dan wirausaha handal; (2) nilai-nilai "*Adek Pangadereng*" menjadikan orang-orang Wajo sangat menghormati, menjunjung tinggi hukum, hak asasi manusia dan pemerintahan yang demokratis; (3) semboyan "*Oreng Madura ta` tako` mateh, tapeh tako` kalaparan`*" telah mengantar orang-orang Madura menjadi perantau dan pekerja keras; (4) sistem *Subak* di Bali tidak hanya menjadikan masyarakat Bali menjadi masyarakat yang rukun dan damai, tetapi juga menjadi masyarakat yang pandai mengatur sistem ekonomi dan pertanian; (5) budaya "sasi" di Maluku, "tara bandu" di Papua atau yang dikenal di Jawa sebagai "pranata mangsa" tidak hanya berperan dalam pelestarian lingkungan, tetapi lebih jauh mampu mempertahankan keselarasan hubungan manusia dengan alam, keselarasan hidup dan pemanfaatan sumberdaya alam secara lebih arif. Dalam hal ini budaya tersebut tidak hanya

menyangkut kearifan ekologis, tetapi juga menyangkut kearifan sosial, politik, budaya, dan ekonomi.

Dalam lingkup nasional karena untuk mengembangkan ujian nasional maka kearifan local yang cocok adalah BHINEKA TUNGGAL IKA. Kata Bhinneka Tunggal Ika itu diambil dari kutipan kitab Sutasoma karangan Mpu Tantular. Semboyan negara ini diambil dari bahasa Jawa kuno. Kata "Bhinneka" artinya beraneka ragam atau berbeda-beda, kata "Tunggal" artinya satu. Sedangkan "Ika" artinya itu. Secara harfiah Bhinneka Tunggal Ika diterjemahkan menjadi "Beraneka Satu Itu", yang bermakna meskipun berbeda-beda tetapi pada hakikatnya bangsa Indonesia tetap satu kesatuan. Semboyan ini memberi rambu-rambu bahwa penyelenggaraan ujian nasional tidak boleh membawa dampak negative terhadap kesatuan dan persatuan bangsa. Mulai dari persiapan, penulisan butir soal, pelaksanaan ujian nasional, monev, dan pelaporan ujian nasional tidak boleh berdampak negative terhadap kesatuan dan persatuan bangsa.

Pada tulisan ini hanya akan difokuskan pada cara menulis soal ujian nasional, namun karena mulai tahun 2019 ujian nasional sudah digantikan dengan Assesmen Kompetensi Minimum (AKM) maka tulisan ini hanya terfokus pada cara penulisan butir-butir soal AKM. Asesmen Kompetensi Minimum (AKM) merupakan penilaian kemampuan minimum yang dilakukan kepada peserta didik. Kemampuan minimum yang dimaksud adalah kemampuan paling dasar yang harus dimiliki oleh peserta didik pada jenjang tertentu, yang meliputi literasi membaca dan numerasi. Kemampuan ini sesuai dengan kecakapan abad ke-21 yang menuntut peserta didik untuk dapat mengikuti perkembangan zaman yang penuh dengan tantangan. Dengan menguasai kecakapan abad ke-21, peserta didik akan memiliki keterampilan belajar dan berinovasi, keterampilan menggunakan dan memanfaatkan teknologi/media informasi, serta dapat bekerja dan bertahan dengan menggunakan kecakapan hidup (*life skill*).

Sementara itu Jones (2018) menjelaskan bahwa pendidikan di abad dua puluh satu memprioritaskan pengembangan individu dan sosial, serta keterampilan yang memadai dilengkapi dengan kemampuan berfikir kritis, kreatif, kemampuan beradaptasi dan kewirausahaan. Hal ini senada dengan pendapat World Economic Forum/WEF (Gleason, 2018) yang menjelaskan "*the top ten skills that will be needed in order of priority by employers by 2020 are: complex problem solving, critical thinking, creativity, people management, coordinating with others, emotional intelligence, judgment and decision making, service orientation, negotiation, and cognitive flexibility*". Di bagian lain World Economic Forum (2015) menetapkan enam literasi dasar, yaitu (a) literasi baca tulis, (b) literasi numerasi, (c) literasi sains, (d) literasi digital, (e) literasi finansial, dan (f) literasi budaya dan kewargaan.

Berdasarkan uraian di atas, maka soal ujian nasional yang disusun harus berkualitas tinggi dilihat dari sisi konstruksi, materi, dan Bahasa. Soal yang disusun harus melatih peserta didik untuk mampu berfikir tingkat tinggi *complex problem solving, critical thinking, dan creativity*. Soal ujian nasional harus memiliki muatan kearifan local, misal tidak mengandung SARA (suku, agama, ras, dan asumsi negative). Agar tidak terjadi kesalahan konsep dan soal ujian nasional ini mampu mendorong pencapaian kecakapan di abad 21 maka Pemerintah telah mengembangkan Framework AKM Tahun 2020. Dalam frame work ini dijelaskan bahwa dalam rangka menyiapkan peserta didik yang memiliki kecakapan abad ke-21, maka asesmen kemampuan minimum (AKM) yang dilakukan meliputi literasi membaca dan numerasi, yaitu:

(a) asesmen pada kemampuan bernalar menggunakan bahasa (literasi membaca) dan (b)

asesmen kemampuan bernalar menggunakan matematika (numerasi).

Literasi membaca bukan hanya sekadar kemampuan membaca secara harfiah tanpa mengetahui isi/makna dari bacaan tersebut, melainkan kemampuan memahami konsep bacaan. Sementara itu, literasi numerasi bukan hanya sekadar kemampuan menghitung, melainkan kemampuan mengaplikasikan konsep hitungan di dalam suatu konteks, baik abstrak maupun nyata. Bentuk soal AKM bervariasi, yaitu: (a) pilihan ganda (PG) 5 pilihan, (b) pilihan ganda kompleks, (c) menjodohkan, (d) isian, dan (e) esai atau uraian.

a. Pilihan Ganda

Soal pilihan ganda terdiri atas pokok soal dengan beberapa pilihan jawaban. Peserta didik diminta menjawab soal dengan memilih satu jawaban benar dari beberapa pilihan jawaban yang disediakan. Jumlah pilihan jawaban untuk soal kelas 1 sampai dengan kelas 3 sebanyak 3 pilihan (A, B, C), kelas 4 sampai dengan kelas 9 sebanyak 4 pilihan (A, B, C, D), dan kelas 10 sampai dengan kelas 12 sebanyak 5 pilihan (A, B, C, D, E). Penulisan soal pilihan ganda harus memenuhi kaidah penulisan soal PG, yaitu dari segi materi, konstruksi, dan bahasa.

- 1) Dari segi materi: sesuai dengan semboyan Bhineka Tunggal Ika, yakni tidak mengandung SARA (suku, agama, ras, dan asumsi negative), melatih peserta didik untuk mampu berfikir tingkat tinggi *complex problem solving*, *critical thinking*, dan *creativity*, konsep harus benar, kunci hanya satu, dan pilihan jawaban harus homogen dan logis.
- 2) Dari segi konstruksi, pokok soal dan pilihan jawaban harus jelas dan tidak menimbulkan pengertian ganda, informasi yang dituliskan hanya yang diperlukan, pilihan jawaban tidak menggunakan kalimat “semua jawaban di atas salah/benar”.
- 3) Dari segi bahasa, soal harus memenuhi kaidah bahasa Indonesia.

b. Pilihan Ganda Kompleks

Soal pilihan ganda kompleks terdiri atas pokok soal dan beberapa pernyataan yang harus dipilih peserta didik dengan memberi tanda centang (✓) pada kotak yang disediakan di depan setiap pernyataan yang dianggap sesuai dengan permasalahan pada pokok soal, pada kolom Ya/Tidak, pada kolom Benar/Salah, atau pilihan lain yang sesuai. Pemberian skor berdasarkan kompleksitas dari pernyataan dan jumlah pilihan jawaban. Apabila jumlah pernyataan 3-5 dan pilihan jawaban 2 (benar-salah, ya-tidak, berubah –tidak berubah, atau lainnya), penskoran 1 atau 0. Artinya, diberi skor 1 bila semua jawaban benar, diberi skor 0 bila ada jawaban salah. Apabila jumlah pernyataan lebih dari 5 dan pilihan jawaban lebih dari 2 (hewan-tumbuhan-mikroorga nisme, pagi-siang-malam, kota-kabupaten-kecamatan-desa, hijau-merah-kuning-biru-oranye, atau lainnya), penskoran 2 1 0. Diberi skor 2 bila menjawab semua benar, diberi skor 1 bila salah 1 atau 2, diberi skor 0 bila salah lebih dari 2.

c. Menjodohkan

Bentuk soal menjodohkan mengukur kemampuan peserta tes dalam mencocokkan, menyesuaikan, dan menghubungkan antardua pernyataan yang disediakan. Soal ini terdiri atas dua lajur. Lajur pertama (sebelah kiri) berupa pokok soal dan lajur kedua (sebelah kanan) berupa jawaban. Jumlah jawaban sebaiknya lebih banyak daripada jumlah pokok soal di sebelah kiri.

d. Isian atau jawaban singkat

Soal isian dan jawaban singkat adalah soal yang menuntut peserta tes untuk memberikan jawaban secara singkat, berupa kata, frasa, angka, atau simbol. Perbedaannya adalah soal isian disusun dalam bentuk kalimat berita, sementara itu soal jawaban singkat disusun dalam bentuk pertanyaan.

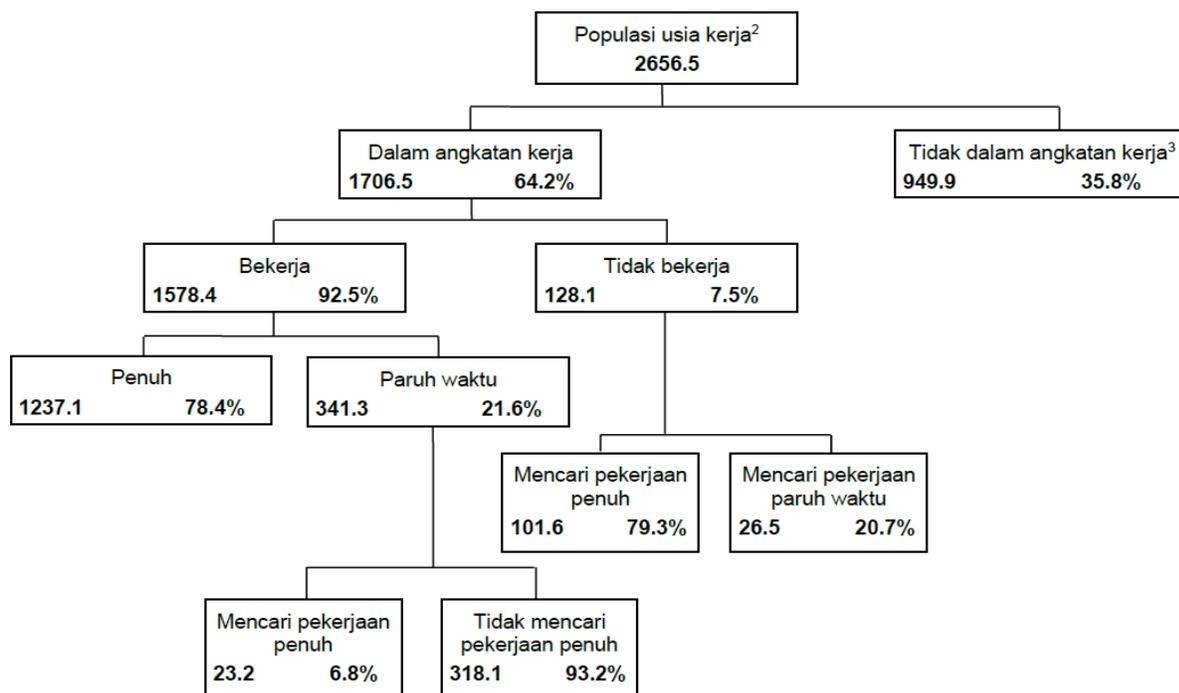
e. Esai atau uraian

Soal uraian adalah soal yang jawabannya menuntut peserta didik untuk mengingat dan mengorganisasikan gagasan-gagasan dengan cara mengemukakan atau mengekspresikan gagasan tersebut dalam bentuk uraian tertulis. Pada soal uraian disediakan pedoman penskoran yang merupakan acuan dalam pemberian skor. Jawaban peserta didik akan diskor berdasarkan kompleksitas jawaban. Skor penuh atau skor tertinggi diberikan pada jawaban yang memenuhi semua kriteria/kunci jawaban benar. Skor sebagian diberikan pada jawaban yang kurang memenuhi kriteria/kunci jawaban benar. Jawaban salah diberi skor 0, sedangkan tidak menjawab atau kosong diberi kode 9. Pemberian skor baik soal pilihan ganda, pilihan ganda kompleks, menjodohkan, maupun isian singkat dilakukan secara objektif. Sementara itu, untuk soal uraian, penskoran dilakukan oleh penskor dengan mengacu pada pedoman penskoran. Pedoman penskoran dibuat oleh penulis soal ketika menulis soal. AKM diadministrasikan menggunakan komputer.

5. Contoh Soal Dari Berbagai Bentuk Soal

Diagram pohon berikut menunjukkan struktur angkatan kerja atau “populasi usia kerja” suatu negara. Jumlah penduduk total negara itu pada tahun 1995 adalah 3,4 juta jiwa.

Tahun Struktur Angkatan Kerja berakhir pada 31 Maret 1995



Catatan :

1. Jumlah orang dinyatakan dalam ribuan (000s).
2. Populasi usia kerja didefinisikan sebagai orang-orang yang berusia antara 15 sampai dengan 65 tahun.
3. Orang-orang yang “Tidak dalam angkatan kerja” adalah mereka yang tidak secara aktif mencari kerja dan/atau tidak dapat bekerja

Gunakan informasi tentang tenaga kerja suatu negara tersebut untuk menjawab pertanyaan-pertanyaan di bawah ini:

a. Contoh Soal Pilihan Ganda

TENAGA KERJA

Populasi usia-kerja dikelompokkan ke dalam dua grup utama yang mana?

A. Bekerja dan tidak bekerja.
B. Usia kerja dan bukan usia kerja.
C. Pekerja penuh-waktu dan pekerja paruh-waktu.
D. Di dalam angkatan kerja dan tidak dalam angkatan kerja.

b. Contoh soal isian

TENAGA KERJA

Berapa orang yang tergolong usia kerja yang tidak berada di dalam angkatan kerja? (Tuliskan **jumlah** orangnya, bukan persentasenya).

.....

c. Contoh soal Pilihan Ganda Komplek

Contoh soal PG Kompleks

TENAGA KERJA

Pada bagian diagram pohon yang mana, jika ada, orang-orang di bawah ini akan dikelompokkan?

Tunjukkan jawabanmu dengan memberi tanda silang pada kotak dalam tabel. Yang pertama telah dikerjakan untuk kamu.

	“Dalam angkatan kerja: bekerja”	“Dalam angkatan kerja: tidak bekerja”	“Tidak dalam angkatan kerja”	“Tidak termasuk kategori apa pun”
Seorang pramusaji paruh waktu, usia 35	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Seorang wanita karier, usia 43, bekerja 60 jam seminggu	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Mahasiswa, usia 21	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Seorang laki-laki, usia 28, yang baru menjual tokonya dan sedang mencari kerja	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Butir soal dianggap pilihan ganda kompleks karena butir soal itu menuntut dua jawaban benar. Pada contoh di atas tampak bahwa kunci butir pertama adalah pilihan pertama karena butir pertama itu menunjukkan bahwa seorang pramusaji itu bekerja dan usia 35 tahun itu termasuk usia angkatan kerja.

TENAGA KERJA

Umpamakan informasi tentang angkatan kerja disajikan dalam diagram pohon seperti ini setiap tahun.

Di bawah ini adalah daftar empat ciri dari diagram pohon. Nyatakan apakah Anda merasa ciri-ciri itu akan berubah dari tahun ke tahun dengan melingkari kata “Berubah” atau “Tidak berubah”. Yang pertama telah dikerjakan sebagai contoh:

Ciri-ciri Diagram Pohon	Jawaban
Label dalam tiap kotak (misal: “Dalam angkatan kerja”)	Berubah / <u>Tidak berubah</u>
Persentase (mis. “64,2%”)	Berubah / Tidak berubah
Jumlah (mis. “2656,5”)	Berubah / Tidak berubah
Catatan kaki di bawah diagram pohon	Berubah / Tidak berubah

Butir pertama, kuncinya adalah TIDAK BERUBAH karena labelnya memang tidak berubah, tetapi isinya yang berubah. Bagaimana butir 2, 3, dan 4?

Contoh soal non-objektif/esai/uraian

TENAGA KERJA

Informasi mengenai struktur angkatan kerja telah disajikan dalam bentuk diagram pohon, tetapi dapat juga disajikan dengan cara lain, seperti deskripsi tertulis, diagram lingkaran, grafik, atau tabel.

Mengapa diagram pohon dipilih dalam menyampaikan informasi tersebut? Jelaskan jawabanmu!

.....

LITERASI MEMBACA

1. Definisi

Literasi membaca adalah kemampuan untuk memahami, menggunakan, mengevaluasi, merefleksikan bentuk-bentuk teks tertulis yang dibutuhkan oleh masyarakat dan/atau dihargai oleh individu. Pembaca dapat membangun makna dari teks dalam berbagai bentuk. Mereka membaca untuk mengembangkan pengetahuan dan potensi serta untuk berpartisipasi dalam masyarakat sebagai warga negara Indonesia dan dunia. Kemampuan individu memahami teks dipengaruhi oleh kecakapan mereka dan kesanggupan mereka mengolah informasi. Kemampuan literasi membaca untuk peserta didik harus ditingkatkan. Dengan kemampuan literasi yang dimiliki, peserta didik dituntut mampu merefleksikan beragam informasi penting yang diperoleh untuk bekal berpartisipasi dalam lingkungan ilmu pengetahuan dan teknologi serta untuk pengembangan kapasitas diri. Selain itu, kemampuan literasi membaca juga diharapkan mampu membentuk karakter, menggali kemampuan berpikir kritis dan kreatif, dan mampu menumbuhkan partisipasi secara positif dalam komunikasi dan kerjasama. Pada era informasi saat ini, aktivitas literasi membaca membutuhkan tingkat berpikir yang lebih tinggi (*higher order thinking*). Perkembangan dunia ilmu pengetahuan membutuhkan kognisi tinggi karena persaingan sosial dan ekonomi yang semakin ketat. Selain itu, peserta didik saat ini berada dalam jalur informasi palsu (*hoax*) yang meluas dan mudah untuk diakses. Oleh karena itu, perlu kemampuan literasi membaca yang memadai agar mampu mengatasi berbagai persoalan sosial dan akademik yang dihadapinya.

Konten teks

Aspek penting dalam pelaksanaan AKM pada literasi membaca adalah ketersediaan teks atau bacaan yang akan digunakan sebagai stimulus dalam penyusunan soal. Teks atau bacaan tersebut harus memenuhi kriteria tingkat keterbacaan yang baik dan berkualitas baik dari segi **konten**, bahasa, maupun penyajiannya. Jika dihubungkan dengan kecakapan hidup abad ke-21, teks atau bacaan yang digunakan dalam AKM harus mampu mengukur sekaligus menumbuhkan kecakapan berpikir kritis dalam pemecahan masalah, berkomunikasi, kreativitas & inovasi, dan berkolaborasi.

Untuk kepentingan penyusunan soal AKM, konten teks dikelompokkan menjadi dua, yaitu teks fiksi dan teks informasi. Melalui teks fiksi peserta didik dapat memperoleh hiburan, menikmati cerita, dan melakukan perenungan untuk menghayati permasalahan kehidupan yang ditawarkan pengarang. Di sisi lain, melalui teks informasi peserta didik dapat memperoleh fakta, data, dan informasi untuk pengembangan wawasan dan ilmu pengetahuan yang bersifat ilmiah.

Contoh teks fiksi yang dapat digunakan sebagai stimulus bacaan dalam penyusunan soal AKM, antara lain cerita rakyat, legenda, fabel, mitos, fiksi ilmiah, satir, puisi, prosa, drama, novel, pantun, soneta, epos, cerita bergambar, cerita fantasi, ironi, lirik lagu, catatan perjalanan, dan biografi/autobiografi.

Contoh teks informasi yang dapat digunakan sebagai stimulus bacaan dalam penyusunan soal AKM, antara lain iklan, dokumen perusahaan/ pemerintahan (nota dinas, undangan, kontrak, pemberitahuan, pengumuman, dan sebagainya), berita, artikel, laporan, pidato, buku pelajaran, pamflet, brosur, buletin, infografis, label (makanan/obat), resep

(makanan/minuman), ulasan (resensi buku/film/drama), jurnal ilmiah, laporan penelitian ilmiah, buku panduan, dan editorial.

Contoh teks informasi

Hampir setahun Masyarakat Ekonomi ASEAN (MEA) diterapkan, namun belum ada perubahan signifikan yang dirasakan masyarakat luas. Segala sesuatunya kembali berjalan normal seperti sediakala, tetapi seperti air tenang yang menghanyutkan. Justru dalam kondisi seperti inilah diperlukan kesungguhan untuk berbenah diri guna mempersiapkan diri menuju MEA.

Menurut Kepala Pusat Studi Sosial Asia Tenggara (PSSAT) UGM, jika pemerintah dan masyarakat tidak segera berbenah diri, dikhawatirkan bangsa Indonesia hanya akan menjadi penonton di negeri sendiri karena tidak sanggup bersaing dengan produk barang dan jasa serta tenaga kerja dari negara tetangga. Sementara itu, menurut Hermin masyarakat Indonesia belum memahami sepenuhnya tentang MEA. Tanpa ada pendampingan dan pemahaman yang diberikan, masyarakat tidak sadar akan perlunya mempersiapkan diri, misalnya mengambil inisiatif untuk menciptakan produk unggulan yang bisa dipasarkan ke seluruh ASEAN. Walaupun punya produk, hal itu hanya untuk dipasarkan di pasar lokal, tanpa memiliki orientasi *go internasional*.

Fakta lain yang menunjukkan Indonesia kalah langkah dari negara ASEAN lain adalah menjelang diberlakukannya MEA banyak mahasiswa Thailand yang berkunjung ke Indonesia untuk mengenal secara mendalam tentang ASEAN. Selain itu, banyak mahasiswa Thailand yang mulai mempelajari bahasa Indonesia, dengan harapan setelah mahir, mahasiswa tersebut bisa bekerja di Indonesia sebagai tenaga profesional. Bahkan, sekolah-sekolah di negara-negara ASEAN (selain Indonesia) sejak dini telah mengenalkan bendera-bendera negara ASEAN kepada para siswa, sedangkan di Indonesia hal itu belum dilakukan. (*Kedaulatan Rakyat*, 16 November 2016)

Berdasarkan bacaan di atas, jawablah pertanyaan-pertanyaan berikut

1. Ide mendasar yang disampaikan melalui wacana tersebut adalah ...
 - A. peringatan pada pemerintah Indonesia agar segera berbenah diri dalam menghadapi MEA
 - B. kekhawatiran dalam menghadapi MEA masyarakat menjadi penonton di negeri sendiri
 - C. tanggapan masyarakat terhadap MEA belum dapat dilihat dalam kehidupan secara nyata
 - D. fakta menunjukkan Indonesia kalah satu langkah dari aktivitas negara ASEAN yang lain

2. Salah satu bukti sebagian besar masyarakat Indonesia belum sepenuhnya memahami MEA adalah ...
 - A. sosialisasi konsep MEA belum dilakukan secara luas dan masif
 - B. kondisi masyarakat kini seperti air tenang yang menghanyutkan
 - C. kerja sama antarnegara ASEAN masih perlu ditingkatkan terus
 - D. produk yang dihasilkan masyarakat masih berorientasi nasional

Distribusi soal AKM berdasarkan konten teks pada setiap jenjang sebagai berikut.

Konten Teks	Kelas 5	Kelas 8	Kelas 11
Teks Informasi	50%	60%	70%
Teks Fiksi	50%	40%	30%

Bahan bacaan literasi AKM dapat mencakup tiga konteks, yaitu:

- (a) konteks personal,
- (b) konteks sosial-budaya, dan
- (c) konteks saintifik.

Konteks Personal

Bahan teks atau bacaan dengan konteks personal adalah teks atau bacaan yang berisi peristiwa, latar, aksi, karakter, atmosfer/suasana, perasaan, ide maupun wawasan yang bersifat personal (individual). Isi bacaan pada konteks personal dapat berupa hobi, cita-cita, peristiwa atau pengalaman pribadinya, memilih/menentukan gaya hidup, pekerjaan/profesi, dan lain-lain yang bersifat personal (individual). Dengan konteks ini diharapkan peserta didik memiliki kemampuan literasi membaca dalam membentuk karakter dengan menggali kemampuan berpikir kritis dan kreatif dalam kehidupan pribadinya.

Berikut ini contoh teks dengan konteks personal.

Si Kikir dan Emasnya
Cerita Rakyat oleh Aesop

Seorang yang kikir menjual seluruh hartanya dan membeli segumpal emas yang dikuburnya di dalam sebuah lubang di samping sebuah dinding tua. Dia kemudian mengunjungi simpanannya itu setiap hari. Salah seorang anak buahnya memperhatikan hal ini dan memutuskan untuk mengintai gerak gerak si kikir. Anak buahnya ini kemudian mengetahui rahasia harta yang tersembunyi tersebut, dan mulai menggali, dan menemukan segumpal emas, dan dicurinya. Si kikir, pada kunjungan berikutnya, menemukan lubang yang sudah kosong dan mulai menarik-narik rambutnya dan meraung-meraung sejadi-jadinya. Seorang tetangga, yang melihat kejadian itu dan mengetahui apa penyebabnya, kemudian berkata, "Berdoalah dan jangan bersedih, ambillah segumpal batu, dan letakkan di dalam lubang itu, dan bayangkan seolah-olah emas itu masih berada di sana. Bagi kamu hal itu akan sama saja, karena sewaktu emas itu berada di sana, kamu tidak memilikinya, karena kamu sedikit pun tidak menggunakannya."

Berikut ini contoh teks dengan konteks sosial-budaya

PENGUMUMAN SUPERMARKET

Peringatan bagi yang Alergi Kacang
Biskuit Krim Lemon

Tanggal pengumuman: 04 Februari
Nama Produsen: Fine Foods Ltd
Informasi Produk: 125g Biskuit Krim Lemon
(Baik digunakan sebelum 18 Juni dan Baik digunakan sebelum 01 Juli)
Keterangan: Beberapa biskuit dalam sejumlah produk ini kemungkinan mengandung potongan kacang, yang tidak tercantum dalam daftar bahan. Mereka yang alergi terhadap kacang disarankan untuk tidak memakan biskuit ini.
Tindakan Konsumen: Bila Anda telah membeli biskuit ini, Anda dapat mengembalikannya dan mendapatkan kembali uang Anda di toko tempat Anda membeli. Atau telepon ke 1 800 034 241 untuk informasi lebih lanjut.

Berikut ini contoh teks dengan konteks saintifik



Konteks	Kelas 5	Kelas 8	Kelas 11
Personal	60%	40%	30%
Sosial-budaya	30%	40%	40%
Saintifik	10%	20%	30%

Level kognitif literasi membaca

Pada literasi membaca AKM, terdapat tiga level kognitif yang diujikan, yaitu:

- (1) menemukan informasi (access and retrieve),
- (2) memahami (interpret and integrate), dan
- (3) mengevaluasi dan merefleksikan (evaluate and reflect)

Level kognitif	Kelas 5	Kelas 8	Kelas 11
Menemukan Informasi (<i>Access and Retrieve</i>)	50%	40%	30%
Memahami (<i>interpret and integrate</i>)	40%	40%	40%
Mengevaluasi dan merefleksi (<i>Evaluate and reflect</i>)	10%	20%	30%

Numerasi

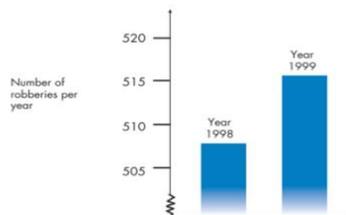
Numerasi adalah kemampuan berpikir menggunakan konsep, prosedur, fakta, dan alat matematika untuk menyelesaikan masalah sehari-hari pada berbagai jenis konteks yang relevan untuk individu sebagai warga negara Indonesia dan dunia. Numerasi dimaknai sebagai kemampuan yang dimiliki oleh seseorang dalam menggunakan pengetahuan matematika yang dimilikinya dalam menjelaskan kejadian, memecahkan masalah, atau mengambil keputusan dalam kehidupan sehari-hari. Hal ini dapat membantu peserta didik mengenali peran matematika dalam kehidupan nyata sehingga dapat membuat penilaian dan keputusan yang diperlukan serta menjadi manusia bertanggung jawab yang mampu bernalar/berpikir logis. Konteks dalam AKM Numerasi mencakup konteks yang dekat dengan dunia peserta didik, sosial, budaya, lingkungan, sains, maupun keilmuan matematika. Konteks-konteks tersebut dikategorikan menjadi tiga, yaitu personal, sosial-budaya, dan saintifik.

Contoh soal dalam konteks personal

Contoh soal konteks sosial budaya:

Seorang reporter berita menunjukkan grafik dan menyampaikan bahwa:

“Grafik menunjukkan bahwa terdapat peningkatan kriminalitas yang sangat pesat dari tahun 1998 ke 1999”



Apakah penafsiran grafik oleh reporter tersebut tepat? Berikan penjelasan.

Adaptasi dari *OECD (2009)*

OECD. 2009. *Take the test. Sample questions from OECD's PISA assessments*. Paris: OECD

Contoh soal dengan konteks saintifik (ekstra-matematika):

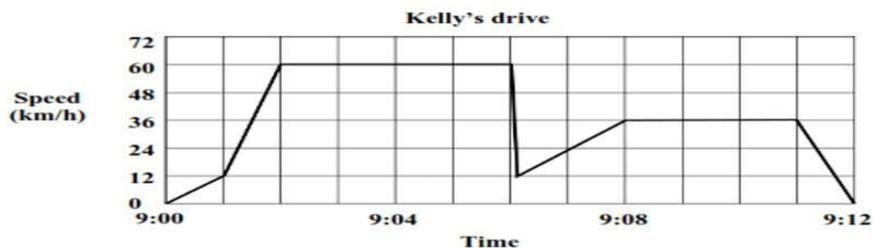
Untuk diobati penyakitnya, seorang pasien di rumah sakit disuntik obat. Tubuh pasien secara bertahap mengolah obat tersebut sehingga setelah 1 jam hanya tersisa 60% obat yang masih aktif. Pola ini berlanjut terus, yaitu di akhir setiap satu jam hanya ada 60% obat dari periode satu jam sebelumnya yang masih aktif.

CAR DRIVE

Kelly went for a drive in her car. During the drive, a cat ran in front of the car. Kelly slammed on the brakes and missed the cat.

Slightly shaken, Kelly decided to return home.

The graph below is a simplified record of the car's speed during the drive.



Question 1: CAR DRIVE

M302Q01

Question intent: Change and relationships

What was the maximum speed of the car during the drive?

Maximum speed: km/h.



Tabel di atas menunjukkan bahwa persentase kontek itu tergantung pada level sekolah. Untuk konteks personal, semakin tinggi tingkat sekolah, semakin kecil persentasenya. Sementara itu untuk kontek saintifik, semakin tinggi level sekolah, semakin besar persentasenya.

SIMPULAN DAN SARAN

Ujian nasional berbasis kelokalan dalam tulisan ini adalah asesmen kompetensi yg berbasis pada BHINEKA TUNGGAL IKA. Soal yg digunakan untuk penilaian dapat berbentuk pilihan ganda, pilihan ganda kompleks, isian, dan uraian. Soal harus berkualitas tinggi dilihat dari aspek materi, konstruksi, dan bahasa. Soal juga harus mendidik peserta tes utk berfikir kritis, memecahkan masalah, dan kreatif.

DAFTAR PUSTAKA

- Anastasi, A., & Urbina, S. (1997). *Psychological testing*. Prentice Hall/Pearson Education.
- Allen, M.J. & Yen, W.M. (1979). *Introduction to measurement theory*. Monterey, CA: Brooks/Cole Publishing Company.
- Bickley, J., Owen, R. J., Fraser, A. G., & Pounder, R. E. (1993). Evaluation of the polymerase chain reaction for detecting the urease C gene of *Helicobacter pylori* in gastric biopsy samples and dental plaque. *Journal of medical microbiology*, 39(5), 338-344.
- Brown, M. B., & Forsythe, A. B. (1974). The small sample behavior of some statistics which test the equality of several means. *Technometrics*, 16(1), 129-132.
- Brown TL, Gavin, Irving, S. Earl, dan Keegan, J. Peter. (2014). *An Introduction to educational assessment, measurement and evaluation: Improving the quality of teacher-based assessment, 3rd Edition*. Auckland: Dunmore Publishing Ltd.
- Cronbach, L. J., & Furby, L. (1970). How we should measure "change": Or should we?. *Psychological bulletin*, 74(1), 68.
- Djemari Mardapi (2017). Pengukuran, penilaian, dan evaluasi pendidikan (edisi 2). Yogyakarta: parama Publishing.
- Dunn, K.E., dan Mulvenon, S.W. (2009). A Critical review of research on formative assessment: The Limited scientific evidence of the impact of formative assessment in education. *Practical Assessment, Research & Evaluation*, 14 (7).
- Fernandes, H.J. X. (1984). *Evaluation of educational program*. Jakarta: National Education Planning, Evaluating and Curriculum Development.
- Fitzpatrick, J. L., Sanders, J. R., & Worthen, B. R. (2004). Program evaluation: Alternative approaches and practical guidelines. Frey, B. B., & Schmitt, V. L. (2007). Coming to terms with classroom assessment. *Journal of Advanced Academics*, 18(3), 402-423. doi:10.4219/jaa-2007-495
- Gronlund, N. E. (1985). *Measurement and evaluation in teaching*. Macmillan College.
- Hargreaves, A., Earl, L., & Schmidt, M. (2002). Perspectives on alternative assessment reform. *American Educational Research Journal*, 39(1), 69-95. doi:10.3102/00028312039001069.
- Jones, P. W. (2018). *International policies for Third World education: UNESCO, literacy and development*. Routledge.
- Kerlinger, F.N. (1979). *Behavioral research: A conceptual approach*. New York: Holt, Rinehart, and Winston.

- Masters, G. N., & Keeves, J. P. (Eds.). (1999). *Advances in measurement in educational research and assessment*. Pergamon.
- Mehrens, W. A. L., & Irvin, J. (1973). *Measurement and Evaluation in Education and Psychology/William A. Mehrens and Irvin J. Lehmann* (No. LB1051. M43 1991.)
- Nunnally, J.C. (1978). *Psychometric theory*. New York: McGraw Hill Book Company. Inc. Permendiknas Nomor 23 Tahun 2016 Tentang Standar Penilaian.
- Permendiknas R.I. Nomor 20 Tahun 2007 tentang Standar Penilaian.
- Popham, W. J. (1995). *Classroom assessment*. Boston: Allyn and Bacon.
- Pusat Asesmen dan Pembelajaran. (2020). *Framework Asesmen Kometensi Minimum*. Jakarta: Badan Penelitian dan Pengembangan dan Perbukuan Kementerian Pendidikan dan Kebudayaan.
- Salkind, N.J. (2013). *Test & measurement for people who hate test and measurement*. Los Angeles: SAGE Publications, Inc.
- Scriven, M. (1967) *The Methodology of Evaluation, dalam Perspective of Curriculum Evaluation*, AERA 1 (ed. Tyler, R. et.al), Chicago : Rand McNally and Company.
- Stiggins, R. and Chapuis, J. (2012). *Introduction to student involved assessment for learning, 2 nd edition*. Boston: Addison Wesley.
- Stone, M. J., Fairbrother, W. J., Palmer III, A. G., Reizer, J., Saier Jr, M. H., & Wright, P. E. (1992). Backbone dynamics of the Bacillus subtilis glucose permease IIA domain determined from nitrogen-15 NMR relaxation measurements. *Biochemistry*, 31(18), 4394-4406.
- Stufflebeam, Daniel L., and Anthony J *Shinkfield*. (2007). *Evaluation Theory, Models, and Applications*. San Francisco: Jossey-Bass.
- Trespeces, F. A. (1993). *The CIPP Model*. Qoezon City: Innotech.
- Umar, J., & Hayat, B. (2000). Efektifitas Pengujian Soal Bentuk Soal Pilihan Ganda dan Benar Salah. *Jakarta: Pusat Pengujian*.
- Worthen, B.R & Sanders, J.R. (1987). *Educational Evaluation, Alternative Approaches and Practical Guideliness*. New York & London: Longman Inc.