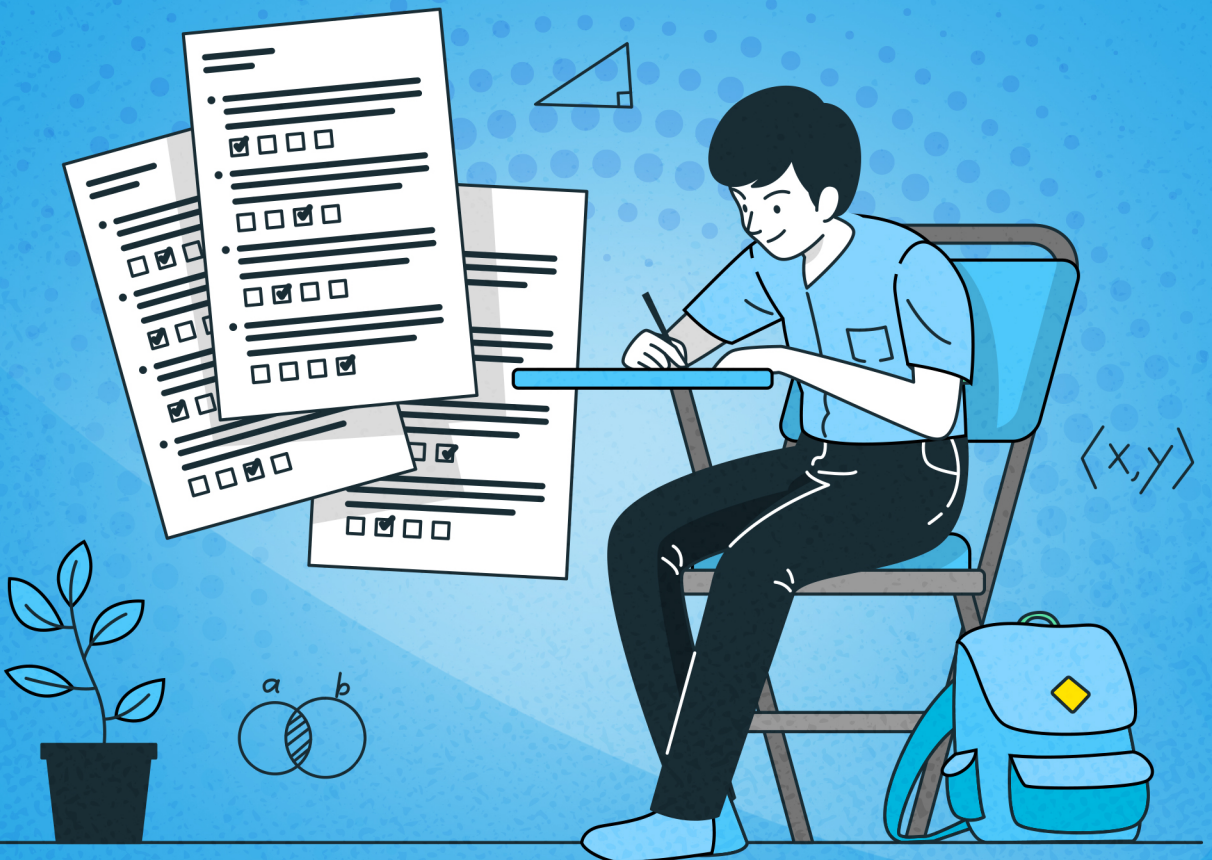


# Kajian

# Akademik & Rekomendasi Reformasi Sistem Asesmen Nasional



# **Kajian Akademik dan Rekomendasi Reformasi Sistem Asesmen Nasional**

Pengusun:

Anindito Aditomo, Ph.D

Dr. Rahmawati

Nisa Felicia, Ph.D

Najeela Shihab, M.Psi

Fiona Handayani, M.B.A

**Badan Penelitian dan Pengembangan dan Perbukuan  
Kementerian Pendidikan dan Kebudayaan  
&  
Pusat Studi Pendidikan dan Kebijakan**

Jakarta, Desember 2019

# Kata Pengantar

Undang-Undang Sistem Pendidikan Nasional (UU Sisdiknas) nomor 20 tahun 2003 menyatakan bahwa pendidikan adalah sebuah usaha sadar dan terencana untuk mewujudkan suasana belajar dan proses pembelajaran agar peserta didik secara aktif mengembangkan potensi dirinya. Proses pembelajaran dan evaluasi hasil belajar merupakan mata rantai yang tidak terputus. Asesmen merupakan bagian dari evaluasi. Setiap asesmen yang diselenggarakan harus memiliki makna dalam perbaikan proses pembelajaran dan sebaliknya ragam aktivitas pembelajaran menuntut penerapan beragam jenis asesmen.

Pemerintah sejak tahun 2003 telah menyelenggarakan Ujian Akhir Nasional (UAN) yang kemudian bertransformasi menjadi Ujian Nasional (UN). UN sebagai alat evaluasi pemerintah terhadap hasil belajar siswa selalu diwarnai pro dan kontra. UN muncul sebagai respon atas kekhawatiran bias mutu hasil belajar antar-satuan pendidikan. Niat baik penyelenggaraan UN sebagai ukuran terstandar untuk mendorong semua pihak berpacu meningkatkan mutu ternyata memiliki sisi negatif. Nilai UN cenderung diposisikan sebagai tujuan utama. Kreativitas dan inovasi penilaian di dalam kelas seringkali terberangus dan mengerucut pada pola penilaian UN.

Pembahasan mengenai peran UN terhadap perbaikan pembelajaran dan perubahan kultur belajar seringkali mendapat perhatian yang kecil. Nilai menjadi tujuan, diagnosa dan umpan balik asesmen terlupakan. Sejatinya asesmen merupakan cermin, alat merefleksi diri. Pemerintah berupaya memberikan cermin yang reliabel dengan migrasi sistem ujian berbasis komputer secara masif dan menyediakan analisis hasil UN secara detail. Namun pemanfaatannya sebagai dasar perbaikan pembelajaran belum optimal. Melakukan asesmen berulang kali tidak akan pernah menghantarkan kepada peningkatan mutu, selama hasilnya tidak dijadikan dasar perbaikan.

Begitu pula dengan isu relevansi UN sebagai alat ukur yang diharapkan mampu mengembangkan potensi unik setiap peserta didik. Meskipun soal-soal kemampuan bernalar mulai menjadi bagian dari UN, tuntutan pengolahan cepat hasil UN berimplikasi kepada keterbatasan format soal yang digunakan. Hal-hal tersebut menghadirkan pertanyaan fundamental mengenai sejauh mana UN mampu mewujudkan tujuan Pendidikan yang diamanahkan UU.

Naskah akademik ini memaparkan sejarah UN sebagai sistem asesmen nasional dan evaluasi Pendidikan dalam UU Sisdiknas. Pemetaan karakteristik setiap asesmen berdasarkan fungsinya ditampilkan untuk menelaah validitas peran UN. Gambaran umum sistem asesmen di beberapa negara serta penelitian terkait asesmen nasional dikupas untuk memberikan wawasan praktik baik bagi Indonesia maupun pengalaman kegagalan yang perlu diwaspadai. Berdasarkan beragam kajian tersebut pada bagian akhir dirumuskan rekomendasi kerangka kebijakan asesmen baru: mengembalikan otoritas penilaian individu siswa kepada pendidik, pemerintah mengevaluasi



sistem bukan individu siswa, pemerintah mengukur kompetensi dasar yang utama, asesmen tidak dilakukan di akhir jenjang untuk optimalisasi perbaikan mutu, serta memotret aspek non-kognitif berupa karakter dan iklim belajar. Sistem asesmen baru didesain berlandaskan prinsip memajukan peserta didik serta meningkatkan mutu hasil belajar secara berkesinambungan.

Akhirnya kami menyampaikan ucapan terima kasih kepada seluruh anggota tim dan semua pihak yang telah berkontribusi dalam penyusunan Kajian Kajian Akademik dan Rekomendasi Reformasi Sistem Asesmen Nasional. Harapan kami, kajian akademik dan rekomendasi reformasi sistem asesmen nasional ini bermanfaat untuk kebijakan yang membawa kebaikan bagi masyarakat, khususnya siswa-siswi Indonesia.

Jakarta, Desember 2019

Kepala Badan Penelitian dan Pengembangan dan Perbukuan

Totok Suprayitno, Ph.D



## Daftar Isi

1. Pengantar .....	1
2. Ujian Nasional sebagai bagian sistem asesmen nasional .....	2
3. Asesmen dan Evaluasi Pendidikan dalam UU Sisdiknas.....	3
4. Reformasi Asesmen Siswa di Manca Negara .....	6
4.1 Amerika Serikat .....	7
4.2 Australia.....	8
4.3 Singapura .....	9
4.4 Hong Kong .....	11
4.5 Cina .....	12
4.6 Belajar dari negara lain: ringkasan .....	12
5. Dampak Tes-tes Terstandar pada Proses Belajar Mengajar.....	13
5.1 Penyempitan kurikulum.....	14
5.2 Distorsi metode mengajar.....	15
5.3 Demoralisasi dan dampak psikologis lain.....	15
5.4 Memicu praktik curang.....	16
5.5 Melanggengkan kesenjangan .....	16
6. Reformasi Sistem Asesmen: Rekomendasi Kebijakan .....	17
6.1 Prinsip reformasi sistem asesmen .....	17
6.2 Rekomendasi sistem asesmen nasional untuk evaluasi sistem .....	20
6.3 Rekomendasi mengenai asesmen untuk kelulusan dan seleksi.....	24
7. Penutup.....	26

## Daftar Tabel

Tabel 1. Dua jenis evaluasi pendidikan yang dimandatkan UU Sisdiknas .....	3
Tabel 2. Karakteristik asesmen yang diperlukan untuk tujuan/penggunaan yang berbeda .....	4
Tabel 3. Komponen dan karakteristik asesmen pengganti UN.....	21
Tabel 4. Perbandingan karakteristik asesmen untuk kelulusan dan untuk seleksi.....	24

# 1. Pengantar

Laporan ini merupakan kajian akademik mengenai sistem asesmen (penilaian) hasil belajar siswa dan perannya dalam meningkatkan mutu sistem pendidikan nasional. Tujuan kajian ini ialah untuk menghasilkan rekomendasi perbaikan kebijakan dan sistem asesmen siswa. Hasil dari asesmen siswa merupakan informasi tentang pencapaian tujuan pembelajaran. Karena itu, asesmen siswa merupakan bagian esensial dari proses belajar-mengajar, serta menjadi landasan utama dari evaluasi program pendidikan.

Kajian reformasi sistem asesmen ini berangkat dari amanah undang-undang mengenai fungsi, tujuan, dan prinsip penyelenggaraan pendidikan (UU no. 20 tahun 2003 tentang Sistem Pendidikan Nasional atau Sisdiknas). Pada tataran sosial, pendidikan berfungsi mengembangkan watak dan peradaban bangsa dalam rangka mencerdaskan kehidupan bangsa (Pasal 3 UU). Pada tataran individual, pasal yang sama memandatkan pendidikan sebagai upaya untuk mengembangkan potensi peserta didik sebagai individu secara utuh, yang mencakup dimensi spiritual (beriman dan bertakwa kepada Tuhan), kognitif (berilmu pengetahuan), afektif-emosional (kreatif dan mandiri), serta sosial (warga negara yang bertanggung jawab). Selanjutnya, UU Sisdiknas menegaskan bahwa "Pendidikan diselenggarakan secara demokratis dan berkeadilan serta tidak diskriminatif dengan menjunjung tinggi hak asasi manusia, nilai keagamaan, nilai kultural, dan kemajemukan bangsa." (Pasal 4).

Kajian ini juga dilandasi oleh prinsip bahwa pendidikan seyogyanya berpihak pada siswa (Shihab, 2019). Reformasi pendidikan kerap, secara implisit maupun terus terang, dipandu oleh perspektif siswa sebagai objek. Misalnya, siswa dipandang sebagai calon pekerja yang harus dibentuk sesuai dengan kebutuhan pasar dan perkembangan ekonomi. Siswa juga kerap dipandang sebagai objek yang harus menerima nilai-nilai kebudayaan yang dianggap penting oleh generasi sebelumnya. Tanpa menyangkal bahwa pendidikan memang perlu melayani fungsi sosial-ekonomi-kultural tersebut, laporan ini mengutamakan perspektif siswa sebagai subjek utama pendidikan. Dalam perspektif ini, tujuan utama reformasi pendidikan – termasuk asesmen – ialah untuk menjadikan sistem lebih efektif dalam menumbuhkan potensi dan kesejahteraan siswa secara utuh.

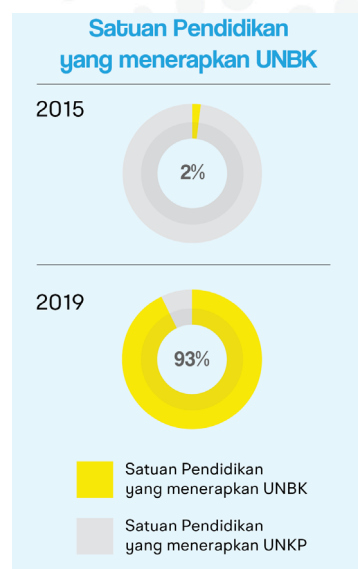
Bagian pertama kajian ini menjelaskan sejarah Ujian Nasional sebagai sistem asesmen nasional yang saat ini ada, sebelum menelaah pasal-pasal UU Sisdiknas yang secara langsung relevan dengan asesmen dan evaluasi pendidikan. Selanjutnya, kajian ini menengok penerapan asesmen di negara-negara lain agar Indonesia dapat belajar dari pengalaman negara lain tanpa perlu mengulang kesalahan-kesalahan yang dilakukan. Bagian selanjutnya merangkum literatur ilmiah yang relevan, untuk memastikan bahwa rekomendasi yang dirumuskan memiliki landasan ilmiah. Bagian akhir kajian ini memaparkan rekomendasi kerangka kebijakan asesmen yang baru.

## 2. Ujian Nasional Sebagai Bagian Sistem Asesmen Nasional

Sejak tahun 1950, sistem ujian nasional telah diterapkan di Indonesia dengan beragam pergantian format serta nama. Dimulai dari ujian penghabisan serta ujian negara yang sangat sentralistis, kemudian masa vakum saat pemerintah hanya memberikan panduan umum ujian satuan pendidikan, digantikan oleh sistem yang memadukan nilai ujian oleh pemerintah (Evaluasi Belajar Tahap Akhir Nasional atau EBTANAS) dan nilai ujian satuan Pendidikan (Evaluasi Belajar Tahap Akhir atau EBTA). Periode EBTA dan EBTANAS tersebut memberikan sinyal kuat adanya praktik pemberian nilai yang tidak objektif oleh guru dan sekolah. Hal ini terlihat dari nilai EBTA yang seringkali nyaris sempurna, namun nilai EBTANAS yang rendah pada siswa yang sama.

Praktik “obral nilai” di sekolah tersebut kemudian memicu lahirnya kebijakan baru: UAN yang didasarkan pada satu standar yang ketat, dengan ancaman ketidakkululusan. Hal ini membuahkan terapi kejut. Pada 2003, ratusan ribu siswa tidak lulus meskipun dengan batas minimal ditetapkan sangat rendah yaitu 3,01 (dalam skala 0-10). Kebijakan satu standar dan sanksi ketidakkululusan berlanjut pada periode UN, dengan batas minimal kelulusan yang ditingkatkan secara bertahap. Kebijakan ini membuat peningkatan hasil UAN/UN menjadi fokus utama bagi *stakeholder* pendidikan. Sayangnya fokus tersebut tidak selalu diterjemahkan menjadi upaya meningkatkan kualitas proses belajar-mengajar. Ada indikasi bahwa tekanan untuk meningkatkan nilai UAN/UN justru mendorong praktik kecurangan secara luas.

Pemerintah berupaya mengatasi praktik kecurangan dengan membuat sistem ujian berbasis komputer yang lebih efisien dan menjamin keamanan bahan tes. UN Berbasis Komputer (UNBK) mulai diterapkan hanya di 400-an satuan pendidikan pada 2015 (2%), dan secara cepat meningkat menjadi lebih dari 89.000 sekolah pada 2019 (93%). Dengan penurunan praktik kecurangan, tingkat kepercayaan terhadap data pun semakin tinggi. Nilai UNBK dapat dianggap lebih merefleksikan penguasaan siswa akan konten kurikulum. Namun tetap tersisa pertanyaan besar: seberapa relevankah kompetensi yang tercermin pada nilai UN dengan tantangan-tantangan Abad 21? Dan seberapa berhasilkah sistem asesmen nasional yang ada saat ini mendorong perubahan positif dalam praktik belajar-mengajar dan mutu hasil belajar siswa?



## 3. Asesmen dan Evaluasi Pendidikan dalam UU Sisdiknas

Dalam konteks Indonesia, reformasi pendidikan tidak bisa dilepaskan dari UU Sisdiknas tahun 2003. UU Sisdiknas tidak secara eksplisit berbicara tentang asesmen siswa. Meski demikian, UU Sisdiknas menyebutkan dua jenis evaluasi pendidikan yang berimplikasi jenis asesmen siswa yang diperlukan. Jenis pertama ialah evaluasi yang bertujuan “memantau proses, kemajuan dan perbaikan hasil belajar peserta didik” (Pasal 58 ayat 1). Evaluasi ini dilakukan oleh guru kepada siswa yang mereka ajar secara berkesinambungan. Jenis kedua ialah evaluasi yang bertujuan “menilai pencapaian standar nasional pendidikan” (Pasal 58 ayat 2). Evaluasi jenis kedua ini dilakukan pada siswa, sekolah, dan program pendidikan oleh lembaga mandiri secara “berkala, menyeluruh, transparan, dan sistemik.” Perbedaan dua jenis evaluasi ini terangkum di Tabel 1.

Tabel 1. Dua jenis evaluasi pendidikan yang dimandatkan UU Sisdiknas

	Evaluasi siswa	Evaluasi sistem
Tujuan utama	Mendorong, memantau kemajuan, dan menilai hasil belajar siswa	Menilai pencapaian standar pendidikan pada level nasional
Pelaksana	Guru	Lembaga mandiri
Prinsip pelaksanaan	Berkesinambungan (sebagai bagian dari proses belajar-mengajar)	Berkala, menyeluruh, transparan, dan sistemik

Pemisahan antara dua jenis evaluasi pendidikan ini secara umum sejalan dengan prinsip dasar asesmen, yakni bahwa asesmen dirancang berdasarkan fungsi atau tujuan penggunaannya (Masters, 2013, 2014). Dalam hal ini, perlu dibedakan antara fungsi evaluasi pada level sistem dan pada level siswa. Sistem pendidikan dapat merujuk pada beberapa tingkat yang berbeda, mulai dari sekolah, dinas pendidikan, sampai nasional. Asesmen hasil belajar siswa menjadi sumber informasi untuk mengevaluasi keberhasilan sistem pendidikan dalam membantu siswa mencapai standar kompetensi yang diharapkan<sup>1</sup>.

<sup>1</sup> Evaluasi sistem terkait dimensi-dimensi lain dari standar pendidikan, seperti pembiayaan, sarana-prasarana, dan pendidik, memerlukan sumber informasi di luar asesmen hasil belajar siswa.



Tabel 2. Karakteristik asesmen yang diperlukan untuk tujuan/penggunaan yang berbeda

	Evaluasi siswa formatif	Evaluasi siswa sumatif	Evaluasi sistem
Tujuan penggunaan hasil asesmen	Mendorong dan memantau kemajuan belajar siswa	Menilai hasil belajar siswa selama periode waktu tertentu	Mendorong, memantau kemajuan, dan menilai efektivitas kebijakan atau intervensi
Target inferensi	Individu siswa	Individu siswa	Kelompok atau populasi
Tingkat kedetilan inferensi	Topik, konsep, atau dimensi kompetensi yang spesifik	Kumpulan topik/konsep dan kompetensi yang lebih menyeluruh	Topik/dimensi spesifik sampai dengan kompetensi yang lebih menyeluruh
Pelaporan hasil asesmen	Informal kepada siswa	Informal dan formal kepada siswa, orangtua, dan sekolah	Formal kepada sekolah, dinas pendidikan, dan publik
Kecepatan pelaporan	Segera ( <i>immediate</i> ) setelah asesmen dilakukan sebagai bagian dari aktivitas belajar-mengajar	Cepat namun tidak harus segera ( <i>immediate</i> )	Bisa lebih lambat
Komparabilitas hasil asesmen	Antarwaktu (secara kualitatif)	Antarwaktu dan antarsiswa (secara kualitatif) di kelas atau sekolah yang sama	Antarwaktu dan antarkelompok/populasi secara kuantitatif
Cakupan domain dan kompetensi	Sempit/spesifik, sesuai keperluan dalam proses belajar-mengajar	Sempit sampai luas, sesuai dengan kurikulum yang digunakan (bisa berbeda antarguru dan sekolah)	Menengah, hanya mencakup domain dan kompetensi yang dipandang perlu dipelajari oleh semua siswa dan sekolah
Sifat instrumen dan skoring	Cenderung tidak terstandar	Tidak harus terstandar, sesuai situasi dan kebutuhan siswa	Harus terstandar
Sifat pelaksanaan	Cenderung tidak terstandar	Bisa namun tidak harus terstandar	Harus terstandar
Pelaksana asesmen	Guru	Guru	Pihak eksternal sekolah

Dari perspektif pemerintah pusat, evaluasi sistem mensyaratkan asesmen berskala besar yang hasilnya bisa digunakan untuk membandingkan skor antarkelompok siswa, antarsekolah, antardaerah, dan antarwaktu secara objektif. Untuk itu, diperlukan instrumen terstandar yang diadministrasikan dan diskor secara terstandar pula. Sebagai konsekuensinya, asesmen seperti ini biasanya dijalankan secara terpusat dengan sistem pelaporan yang terjadwal. Sistem pengolahan terpusat tersebut – pada kasus Ujian Nasional, ditambahkan dengan kebutuhan cepat agar nilai dapat digunakan untuk proses seleksi masuk sekolah – berdampak pada pilihan proses yang lebih mudah dan hemat waktu. Instrumen didominasi oleh soal-soal yang dapat diskor secara otomatis (seperti soal pilihan ganda). Yang juga penting digarisbawahi ialah asesmen untuk evaluasi sistem tidak harus dilakukan untuk tiap siswa, karena inferensi atau kesimpulan yang hendak diambil ada pada level populasi (kelompok siswa di sebuah sekolah, daerah, dst). Karakteristik lain asesmen untuk evaluasi sistem ditampilkan di Tabel 2 (kolom paling kanan).

Evaluasi siswa memerlukan asesmen dengan karakteristik yang berbeda dari asesmen skala besar. UU Sisdiknas menyebutkan beberapa fungsi evaluasi siswa. Fungsi tersebut dapat dikategorikan menjadi dua: formatif dan sumatif (lihat kolom kedua dan ketiga pada Tabel 2). Evaluasi lebih bersifat formatif ketika dimaksudkan untuk memfasilitasi kemajuan belajar (prospektif). Evaluasi lebih bersifat sumatif ketika hasilnya digunakan untuk menilai pencapaian siswa setelah menjalani proses dalam periode waktu tertentu (*retrospective*).

Jika dimaksudkan untuk memfasilitasi kemajuan belajar, maka hasil asesmen perlu segera diketahui dan dikomunikasikan pada siswa. Karena tiap siswa memiliki titik awal dan kecepatan belajar yang berbeda, instrumen dan cara pelaksanaan asesmen yang sesuai juga bisa berbeda untuk tiap siswa. Dengan kata lain, asesmen seperti ini tidak selalu cocok dilakukan dengan instrumen yang terstandar. Pelaksanaannya juga tidak harus terstandar dan formal. Apa yang dinilai juga cenderung spesifik: pemahaman konsep atau keterampilan spesifik, sesuai dengan apa yang sedang dipelajari oleh siswa.

Jika dimaksudkan sebagai penilaian sumatif, asesmen perlu dirancang untuk mencakup semua bagian dari kurikulum. Meski demikian, ini tidak berarti asesmen harus dilakukan secara terstandar. Kompetensi yang sama bisa ditunjukkan dengan cara yang berbeda antarsiswa. Misalnya, sebagian siswa mungkin yang lebih nyaman menunjukkan pemahamannya melalui tanya jawab atau melalui representasi visual semacam *mind map* daripada tes tertulis. Selain itu, kurikulum satu sekolah belum tentu sepenuhnya sama dengan kurikulum sekolah lain. Variasi antarindividu dan antarsekolah ini sulit ditangkap oleh tes terstandar, apalagi jika pengadministrasian dilakukan terstandar secara nasional.

Uraian ini menunjukkan bahwa – sejalan dengan UU Sisdiknas – evaluasi siswa memang seharusnya menjadi ranah kerja dan tanggung jawab guru. Hanya guru yang memiliki akses untuk melakukan asesmen guna mendiagnosis kebutuhan belajar siswa secara individual.



*Hanya guru yang dapat mengumpulkan informasi yang cukup komprehensif untuk menilai pencapaian belajar siswa selama satu periode waktu.*

Tantangan bagi Indonesia ialah memastikan guru-guru kita memiliki kompetensi penilaian yang memadai untuk mampu memilih model penilaian yang tepat, menyusun instrumen penilaian, mengolah hasil, kemudian memaknai hasil penilaian tersebut sebagai informasi untuk perbaikan pembelajaran. Sebuah tes terstandar dari pihak eksternal (seperti UN dan USBN) tidak akan mampu memberikan diagnosa kondisi siswa secara utuh (kognitif, afektif, karakter dan psikomotorik) dan memberikan umpan balik perbaikan pembelajaran yang disesuaikan (*tailored*) terhadap kondisi individu siswa.

Ringkas kata, evaluasi sistem dan evaluasi siswa memerlukan sistem asesmen yang berbeda. Saat ini UN memang sudah tidak lagi menjadi syarat kelulusan, namun masih digunakan sebagai instrumen untuk mengevaluasi individu siswa. UN digunakan juga sebagai alat proses seleksi masuk jenjang pendidikan yang lebih tinggi. Hal ini tentunya kurang tepat, namun tidak dapat dipungkiri alat seleksi diperlukan. Maka menyikapi hal tersebut, perlu dilakukan pemetaan sistem penilaian secara utuh untuk mengakomodir beragam fungsi: diagnosis perbaikan pembelajaran, pemantauan mutu, seleksi, kelulusan (pemberian kredensial), ataupun *benchmarking*. Asesmen perlu ditempatkan kembali (*re-positioning*) sesuai fungsinya, agar tidak lagi terjadi pencampuran fungsi penilaian antara evaluasi sistem dan individu siswa.

Sebelum memberi rekomendasi sistem asesmen yang komprehensif untuk menjawab kebutuhan beragam fungsi penilaian, laporan ini mengulas pengalaman reformasi kebijakan dan sistem asesmen di beberapa negara.

## 4. Reformasi Asesmen Siswa di Manca Negara

Reformasi asesmen siswa sering digunakan oleh pemerintah di berbagai negara untuk mendorong terjadinya perubahan yang lebih luas dalam sistem pendidikan (Supovitz, 2009). Secara garis besar terdapat dua arah yang berbeda terkait reformasi asesmen ini. Yang pertama ialah reformasi yang mengarah pada penggunaan tes-tes terstandar sebagai instrumen untuk mengevaluasi – atau setidaknya memonitor – kinerja guru dan sekolah. Contoh negara yang

menerapkan reformasi semacam ini yaitu Amerika Serikat dan Australia. Yang kedua ialah reformasi yang berorientasi pada penguatan praktik asesmen kelas yang lebih komprehensif dan memiliki fungsi formatif bagi siswa dan guru. Reformasi semacam ini diterapkan oleh negara-negara berprestasi tinggi di Asia seperti Singapura, Hong Kong, Cina, Korea, Taiwan, dan Jepang (Zhao, 2015).

## 4.1 Amerika Serikat

Reformasi asesmen siswa di Amerika terkait erat dengan undang-undang *No Child Left Behind* (NCLB) yang disahkan pada 2002 (Hoffmann, 2003). NCLB dilatarbelakangi oleh hasil kajian *National Commission on Excellence in Education* yang terbit sebagai laporan berjudul “*A Nation at Risk*” yang menyoroti rendahnya mutu sekolah di Amerika dibanding negara-negara maju lainnya, serta besarnya kesenjangan hasil belajar antarkelompok ras dan sosial-ekonomi. Laporan tersebut menyebutkan, misalnya, bahwa 13% dari siswa berusia 17 tahun di Amerika dapat dikatakan buta huruf fungsional. Di kalangan siswa minoritas, angka buta huruf fungsional bahkan setinggi 40%, alias hampir 4 kali lipat dibanding rata-rata nasional (Hoffmann, 2003).



Untuk meningkatkan mutu dan mengurangi kesenjangan, NCLB meminta tiap negara bagian untuk menerapkan tes terstandar matematika dan bahasa Inggris pada setiap siswa kelas 3 sampai 8 (serta satu kali di jenjang sekolah menengah atas). Tes-tes tersebut diharuskan mengacu pada standar pencapaian yang dipandang cukup tinggi atau ambisius untuk mendorong peningkatan mutu. Hasil tes harus dilaporkan untuk siswa secara keseluruhan di tiap sekolah, serta untuk

siswa dari kelompok tertentu yang dianggap tertinggal, seperti minoritas ras/etnis, keluarga miskin, serta siswa berkebutuhan khusus. Tiap negara bagian menetapkan target persentase siswa yang memenuhi standar minimum pencapaian matematika dan bahasa Inggris (yang disebut *Adequate Yearly Progress*) yang harus dipenuhi oleh sekolah. Sekolah dipandang berhasil ketika ada peningkatan skor tes dari tahun ke tahun. Sekolah yang gagal meningkatkan skor tes siswa terancam mendapat berbagai sanksi, mulai dari pengurangan kewenangan pengelolaan dana bantuan sampai penutupan sekolah (Klein, 2015).

Uraian ini menunjukkan bahwa penggunaan tes terstandar secara *high stakes* menjadi elemen kunci dari strategi NCLB untuk meningkatkan mutu dan mengurangi kesenjangan. Seberapa berhasilkah kebijakan tersebut? Beberapa studi menunjukkan bahwa kebijakan

NCLB memiliki dampak positif namun kecil pada prestasi, terutama untuk matematika dan terutama di kelas-kelas awal sekolah dasar (Dee & Jacob, 2011; Hanushek & Raymond, 2005). Meski demikian, NCLB tampaknya tidak berhasil mengurangi kesenjangan pendidikan antarkelompok etnis dan kelas sosial-ekonomi (Huntington-klein & Ackert, 2018). Telaah ulang menggunakan data yang lebih komprehensif menunjukkan bahwa kalau pun NCLB membawa dampak positif, hal itu lebih karena sebagian distrik terdorong untuk menguatkan kapasitas guru dan sumber daya sekolah, bukan semata-mata akibat tekanan dari tes-tes terstandar (J. Lee & Reeves, 2012).

Di luar dampaknya pada prestasi siswa, NCLB juga memiliki beragam dampak lain yang tidak semua bersifat positif. Salah satunya, tes-tes terstandar memicu terjadinya penyempitan kurikulum, yakni berkurangnya kesempatan belajar di luar pelajaran matematika dan bahasa Inggris (Berliner, 2011). Kesempatan pengembangan profesional bagi guru di luar pengajaran matematika dan bahasa juga menjadi lebih sempit (Smith & Kovacs, 2011). Hal ini dapat dilihat sebagai respon yang rasional: dengan sumber daya yang terbatas, guru, sekolah, dan dinas pendidikan akan memilih menggunakan sumber daya tersebut pada hal-hal yang menjadi kriteria evaluasi dalam NCLB. Bahkan untuk pelajaran matematika dan bahasa Inggris pun, NCLB mendorong guru untuk menekankan latihan soal dan persiapan tes tertulis (Erskine, 2014; Schoen & Fusarelli, 2008). Karena itu, dampak positif pada prestasi matematika dasar perlu dilihat dalam konteks berkurangnya kesempatan belajar di luar apa yang secara langsung diukur oleh tes-tes terstandar.

## 4.2 Australia

Reformasi asesmen di Australia menjadi salah satu komponen utama paket kebijakan “revolusi pendidikan” yang digagas oleh pemerintah federal mulai 2007. Pada 2008, pemerintah federal meluncurkan *National Assessment Program - Literacy and Numeracy* (NAPLAN) yang mengukur kecakapan literasi dan numerasi siswa kelas 3, 5, 7, dan 9 di semua sekolah di Australia. NAPLAN dimaksudkan sebagai asesmen yang bersifat *low-stakes* bagi siswa. NAPLAN juga diharapkan menjadi sumber informasi untuk mengidentifikasi area kekuatan dan kelemahan di tiap sekolah, serta merefleksikan dan merencanakan aksi perbaikan (Gable & Lingard, 2013).

Mulai 2010, hasil NAPLAN ditampilkan secara terbuka melalui website *MySchool*. Website tersebut menampilkan rata-rata skor literasi dan numerasi, disertai dengan *Indeks of Community Socio-Educational Advantage* (ICSEA), setiap sekolah di Australia (yang berjumlah sekitar 9500). ICSEA dihitung berdasarkan status sosial-ekonomi siswa yang

dilayani oleh masing-masing sekolah. Informasi tersebut dimaksudkan untuk mendorong publik menakar kinerja sekolah secara lebih “adil” (dengan membandingkan antara sekolah dengan ICSEA yang setara). Publikasi hasil ini menegaskan fungsi NAPLAN sebagai instrumen untuk menilai atau mengevaluasi kinerja sekolah secara publik (Gable & Lingard, 2013). Dengan kata lain, fungsi NAPLAN bergeser atau bertambah: bukan hanya untuk diagnosis dan refleksi, namun juga akuntabilitas.

## NAPLAN

NATIONAL ASSESSMENT PROGRAM  
Literacy and Numeracy

Berbagai studi menunjukkan bahwa meski dimaksudkan sebagai asesmen *low stakes*, NAPLAN memicu perubahan kebijakan akademik sekolah dan cara guru mengajar (Polesel, Rice, & Dulfer, 2013; Swain, Pendergast, & Cumming, 2018; Thompson, 2013). Sebuah survei nasional pada hampir 8000 guru, misalnya, menunjukkan bahwa NAPLAN mendorong pengurangan waktu dan nilai penting pelajaran yang dianggap tidak diuji (Polesel et al., 2013). Sebagian besar guru dalam survei tersebut juga merasakan tekanan untuk mengajar dengan cara latihan ujian (*teaching to the test*), serta mengurangi penerapan strategi pengajaran lain. Guru-guru tersebut mempersepsi NAPLAN lebih sebagai instrumen pemeringkatan dan pengawasan sekolah ketimbang alat diagnosis yang membantu guru dalam mengajar (Polesel et al., 2013).

Siswa cenderung bingung dengan tujuan NAPLAN serta melihatnya sebagai tes yang menekan, terutama jika sekolahnya memberi perhatian ekstra pada persiapan tes (Howell, 2017; Swain et al., 2018). Mereka cenderung tidak menyukai NAPLAN dan merasa bahwa asesmen kelas yang dikelola oleh guru lebih bermakna (Wyn, Turnbull, & Grimshaw, 2014). Sebagian kecil siswa melaporkan gejala kecemasan yang lebih parah seperti insomnia, sakit perut, dan sakit kepala terkait persiapan NAPLAN (Wyn et al., 2014).

Dampak NAPLAN terhadap hasil belajar siswa Australia dapat ditakar dari tren hasil tes internasional seperti PISA, maupun NAPLAN sendiri. Tren skor NAPLAN dalam periode 2010–2019 menunjukkan adanya peningkatan kemampuan membaca dan numerasi siswa sekolah dasar dan menengah. Di sisi lain, data NAPLAN juga menunjukkan bahwa kemampuan menulis menjadi stagnan atau justru menurun. Melengkapi potret ini, data PISA menunjukkan bahwa literasi sains, matematika, maupun membaca siswa sekolah menengah di Australia merosot secara signifikan dalam periode 2000–2018 (Goss, 2019).

### 4.3 Singapura

Singapura secara konsisten menduduki peringkat atas dalam studi-studi asesmen hasil belajar internasional seperti PISA dan TIMSS. Meski demikian, Singapura secara

sadar terus berusaha memperbaiki sistem pendidikannya, termasuk dalam hal praktik asesmen siswa. Era reformasi pendidikan Singapura saat ini dilandasi oleh visi *Thinking Schools Learning Nations* (TSLN) yang diluncurkan sekitar tahun 1997. Fokus TSLN yaitu pada disposisi belajar sepanjang hayat dan kompetensi berpikir kritis dan kreatif, serta berkomunikasi dan bekerja sama (Gopinathan, 1999). Visi TSLN mendapat formulasi yang operasional dan populer dalam seruan *Teach Less, Learn More* (H. L. Lee, 2004).

Dalam hal asesmen, visi TSLN adalah mendorong untuk mempraktikkan asesmen yang lebih otentik dan berorientasi pada proses belajar (Tan, 2011). Sebagai contoh, karya hasil *project* kelompok mulai digunakan sebagai bagian dari kriteria seleksi masuk universitas. Ujian *high stakes* di level pra-universitas juga mulai melibatkan asesmen unjuk kinerja (*performance assessment*) seperti praktikum melakukan penelitian di bidang sains, penggunaan sumber-sumber primer di bidang sejarah, analisis studi kasus di bidang ekonomi, dan merancang karya seni dan aplikasi komputer (Sellen, Chong, & Tay, 2006). *Performance assessment* semacam ini tidak bisa dilakukan pada akhir tahun ajaran, dalam waktu satu atau dua jam sebagaimana ujian konvensional.



Seberapa efektif reformasi asesmen ini mengubah praktik belajar-mengajar di kelas sesuai dengan visi TSLN? Sebuah penelitian mengumpulkan berbagai tugas dan hasil kerja siswa dalam proses belajar-mengajar sehari-hari dalam berbagai pelajaran di 30 sekolah dasar dan 29 sekolah menengah (Koh & Luke, 2009). Data ini mewakili praktik asesmen kelas yang dipraktikkan sekitar 8 sampai 9 tahun setelah TSLN dicanangkan. Analisis terhadap ribuan tugas dan karya siswa dalam penelitian tersebut menunjukkan bahwa praktik asesmen di sebagian besar sekolah yang diteliti masih berorientasi pada proses berpikir tingkat rendah terkait pengetahuan faktual dan prosedural (Koh & Luke, 2009).

Menurut Tan (2011), reformasi asesmen di Singapura terhambat oleh kuatnya budaya *quality assurance* dan penggunaan ujian terstandar untuk menentukan karir akademik siswa. Penekanan *quality assurance* pada indikator yang mudah dikuantifikasi tidak sejalan dengan asesmen otentik yang bersifat kualitatif dan holistik (Tan, 2011). Selain itu, penggunaan ujian terstandar meneguhkan konsepsi asesmen sebagai penilaian yang seharusnya dilakukan secara individual dan dalam kondisi terkontrol (dengan pengawasan dan waktu yang singkat) (Tan, 2011). Konsepsi ini cenderung menyulitkan adopsi bentuk-bentuk asesmen berbasis *project* yang bersifat kolaboratif dan bersifat interaktif serta berjangka waktu panjang.

## 4.4 Hong Kong

Sebagaimana siswa Singapura, siswa Hong Kong juga merupakan juara internasional dalam hasil belajar. Sistem pendidikan Hong Kong juga menekankan penggunaan asesmen sebagai penilaian sumatif untuk menyeleksi dan menentukan karir akademik siswa (Biggs, 1996). Dua gelombang reformasi pendidikan berupaya agar asesmen lebih banyak digunakan untuk mendukung proses belajar-mengajar (Berry, 2011).

Gelombang reformasi yang pertama, yang kerap disebut sebagai inisiatif *Target-Oriented Curriculum Reform* (TOC), menekankan pentingnya asesmen untuk mengetahui level atau perkembangan siswa dalam perjalanannya mencapai standar kompetensi yang ditetapkan kurikulum (Morris, 2002). Dalam inisiatif yang berlangsung pada tahun 1990-an ini, guru diminta mencatat hasil belajar siswa secara terperinci. Banyak guru yang merasa tuntutan tersebut terlalu sulit dan terlalu memakan waktu. Karena luasnya resistensi guru, orang tua, dan sekolah (Carless & Carless, 2005), TOC dinilai sebagai inisiatif yang kurang berhasil mengubah praktik asesmen secara signifikan.

Gelombang reformasi pendidikan kedua, yang dimulai pada awal 2000-an, secara eksplisit memberi penekanan pada *Assessment for Learning* (AfL), alias penggunaan asesmen untuk mendukung proses belajar mengajar (Berry, 2011). Dalam dokumen kurikulum, asesmen dinyatakan sebagai bagian tak terpisahkan dari proses pengajaran, dengan penekanan pada pengumpulan informasi untuk memahami siswa dan memberi umpan balik pada siswa, guru, sekolah, dan orang tua (*Curriculum Development Council* (CDC), 2002). Salah satu komponen utama inisiatif AfL ini yaitu *Basic Competency Assessment* (BCA), seperangkat sumber daya daring (*online*) yang dirancang untuk membantu guru sekolah dasar dalam membuat tugas-tugas asesmen yang tepat untuk kelasnya (Berry, 2011). Komponen lainnya yaitu *School-Based Assessment* (SBA), sekumpulan instrumen asesmen untuk digunakan oleh sekolah dan hasilnya bisa menjadi bagian dari penilaian akhir jenjang sekolah menengah atas (Berry, 2011).

Meski lebih komprehensif dan lebih banyak didukung oleh sekolah dibanding inisiatif TOC, inisiatif AfL juga dinilai kurang berhasil mengubah praktik asesmen di Hong Kong. Banyak sekolah di Hong Kong yang masih menekankan penggunaan asesmen secara sumatif sebagai penilaian akhir bagi prestasi siswa (Berry, 2011). Salah satu faktor utama yang ditengarai menjadi penghambat reformasi asesmen di Hong Kong ialah masih dominannya peran ujian eksternal yang bersifat *high stakes* sebagai alat seleksi untuk memasuki jenjang pendidikan selanjutnya (Kennedy, Chan, Fok, & Yu, 2008).



## 4.5 Cina

Menjelang awal abad 21, Cina menerapkan reformasi pendidikan yang berorientasi pada pengembangan kreativitas dan kemampuan menyelesaikan masalah (Zhao, 2015). Orientasi ini memperluas konsepsi tradisional masyarakat Cina yang melihat bahwa tugas utama sekolah ialah untuk membantu siswa menguasai pengetahuan akademik. Dengan orientasi baru, guru dan sekolah di Cina diharapkan lebih menekankan pada pengembangan karakter pembelajar sepanjang hayat, olahraga, seni, vokasi, dan tanggung jawab sosial.

Reformasi ini tercermin dalam perubahan kurikulum nasional. Pada tahun 2001, jumlah jam pelajaran dalam kurikulum nasional kelas 1-9 berkurang 380 jam. Jam pelajaran matematika berkurang 146 jam, sedangkan jam pelajaran olahraga bertambah sebanyak 156 jam. Untuk sekolah menengah atas, kurikulum yang baru memberi porsi lebih besar pada mata pelajaran pilihan untuk memberi ruang bagi pengembangan minat individual siswa. Sebagai konsekuensinya, mata pelajaran wajib dikurangi jam pelajarannya sebanyak 347 jam (Zhao, 2015).

Dalam hal asesmen, Cina melarang penggunaan tes terstandar untuk siswa kelas 1 sampai 3 sekolah dasar. Untuk kelas 4 dan seterusnya, tes terstandar hanya boleh dilakukan maksimal sekali setiap semester untuk pelajaran tertentu saja seperti Bahasa Cina dan matematika. Tes jenis lain juga dibatasi. Secara umum, siswa tidak boleh diminta menjalani ujian lebih dari dua kali tiap semester (Zhao, 2015). Distrik Shanghai melangkah lebih jauh lagi dengan mengembangkan sistem asesmen yang komprehensif. Penilaian siswa tidak dilakukan hanya berdasarkan ujian atau tes terstandar pada akhir tahun ajaran, melainkan berdasarkan portofolio yang mencakup skor tes, hasil tugas-tugas di tiap pelajaran, aktivitas dan pencapaian personal di luar pelajaran, serta aktivitas sosial di luar sekolah. Shanghai juga mengurangi bobot tes akademik dalam seleksi masuk perguruan tinggi (Zhao, 2015).

## 4.6 Belajar dari Negara Lain: Ringkasan

Uraian di atas memberi ilustrasi dua arah reformasi asesmen. Reformasi jenis pertama mengandalkan tes terstandar sebagai instrumen “akuntabilitas vertikal” di mana guru dan sekolah dievaluasi oleh pemerintah berdasarkan skor ujian siswa mereka. Reformasi jenis kedua secara sistematis mengurangi peran tes terstandar dan menguatkan praktik asesmen kelas untuk mendukung pengembangan kompetensi yang lebih holistik.



Dua arah reformasi asesmen ini memberi pelajaran tersendiri. Reformasi asesmen di Singapura, Hong Kong, dan Cina didorong oleh kesadaran bahwa kuatnya pengaruh ujian tertulis dapat menghambat praktik belajar-mengajar yang berorientasi pada kompetensi Abad 21 yang kompleks. Karena itu, meski sistem pendidikannya tergolong sebagai yang paling efektif di dunia, Singapura, Hong Kong dan Cina berupaya mengurangi pengaruh ujian tertulis yang terstandar, dan mendorong praktik asesmen kelas yang lebih inovatif dan komprehensif seperti *performance assessment*. Upaya reformasi ini mendapat hambatan serius dari konsepsi dan budaya yang telah berakar lama mengenai asesmen dan pengajaran.

Di sisi lain, negara seperti Amerika dan Australia berupaya menggunakan ujian terstandar sebagai instrumen kebijakan untuk mencapai sasaran reformasi pendidikan yang lebih luas. Di Amerika, tes terstandar menjadi alat untuk mengevaluasi mutu sekolah, dengan konsekuensi besar bagi sekolah yang dinilai gagal. Di Australia, NAPLAN digunakan untuk menguatkan standar proses dan hasil belajar yang tadinya bervariasi antarnegara bagian. Meski tidak dimaksudkan sebagai *high stakes* bagi siswa, baik NCLB maupun NAPLAN ternyata memiliki dampak yang tidak selalu positif terhadap praktik belajar-mengajar di kelas. Dampak tes terstandar terhadap praktik pengajaran ini dibahas lebih mendalam di bagian berikut.

## 5. Dampak Tes-tes Terstandar pada Proses Belajar Mengajar

Jika asesmen untuk evaluasi sistem menjadi tanggung jawab pemerintah, asesmen kelas merupakan ranah dan tanggung jawab guru dan sekolah. Meski demikian, perlu diingat bahwa asesmen terstandar yang diselenggarakan pemerintah dapat berpengaruh besar terhadap praktik pengajaran guru di kelas. Karena itu rancangan asesmen untuk evaluasi sistem perlu memperhatikan potensi dampak, baik yang diinginkan maupun yang tidak.

Asesmen diyakini secara luas sebagai instrumen yang efektif untuk mengarahkan perubahan atau reformasi (Looney, 2014). Asesmen nasional cenderung dipersepsi memiliki kredibilitas dan nilai sosial yang lebih tinggi dibanding asesmen lokal atau kelas. Karena itu, asesmen nasional dapat mendorong siswa untuk lebih sering hadir di kelas (dengan kata lain, mengurangi absensi) dan memotivasi perilaku belajar. Bagi guru, asesmen nasional juga dapat menjadi model dalam

hal cara mengevaluasi kompetensi siswa. Asesmen nasional juga dapat memberi sinyal tentang pengetahuan dan keterampilan yang penting dikembangkan (Gregory & Clarke, 2010).

Menurut UNESCO Institute for Statistics (2018), asesmen skala besar memiliki beragam manfaat seperti menjadi landasan empiris bagi revisi kebijakan untuk memperbaiki sistem pendidikan, termasuk untuk afirmasi dan pemerataan kualitas, perbaikan proses belajar-mengajar, perbaikan kurikulum nasional, monitoring dan akuntabilitas, dan perbandingan atau *benchmarking* antarwaktu atau dengan sistem pendidikan negara lain. Namun terwujud tidaknya berbagai manfaat ini tergantung pada bagaimana data hasil asesmen dimanfaatkan dan dijadikan dasar pengambilan kebijakan yang baik. Selain itu, perlu digarisbawahi bahwa data untuk evaluasi sistem tidak harus berupa data individu siswa.

Sayangnya, beberapa literatur ilmiah juga menunjukkan bahwa asesmen nasional berdampak negatif pada sistem pendidikan. Berikut ringkasan dampak-dampak tersebut (Emler, Zhao, & Wang, 2019)

### Dampak Tes-tes Terstandar pada Proses Belajar Mengajar



Penyempitan kurikulum



Distorsi metode mengajar



Memicu praktik curang



Melanggengkan kesenjangan



Demoralisasi dan dampak psikologis lain

## 5.1 Penyempitan kurikulum

Pelajaran yang diukur oleh asesmen yang diselenggarakan pemerintah rentan dipandang lebih penting dibandingkan pelajaran lain. Hal ini dapat menyebabkan penyempitan kurikulum. Sebagai contoh, di Amerika tes bahasa dan matematika digunakan di level negara bagian untuk mengukur keberhasilan/kegagalan sekolah. Hal ini mendorong sekolah menambah antara 75 sampai 150 menit per minggu untuk pelajaran bahasa dan matematika, dengan mengurangi jatah waktu pelajaran ilmu sosial, sains, olahraga, seni, dan musik (Berliner, 2011). Di Cina, siswa dan guru mengalokasikan lebih banyak waktu belajar-mengajar untuk pelajaran yang diujikan dalam seleksi perguruan tinggi nasional (Yu & Suen, 2005). Untuk konteks Indonesia, bukti-bukti anekdot menunjukkan bahwa sekolah menambah waktu instruksional bagi pelajaran yang diujikan secara nasional (KSD,

2016; Puslitbang Sindo, 2018). Selain itu, penyempitan juga terjadi pada kesempatan belajar guru pelajaran non-UN. Riset menemukan bahwa UN menciptakan “kasta” guru: mereka yang mengajar pelajaran UN dipandang lebih penting dan lebih sering diberi kesempatan mengikuti pelatihan dibanding guru lain (Ashadi & Rice, 2016).

## 5.2 Distorsi metode mengajar

Asesmen nasional dapat mendorong penggunaan metode-metode mengajar yang berorientasi sempit pada pengerjaan soal tes (Towles-reeves, Garrett, & Burdette, 2006). Contoh metode yang sempit yaitu pembahasan soal-soal tes tahun sebelumnya dan simulasi-simulasi ujian. Dalam konteks Indonesia, siswa yang memiliki sumber daya finansial seringkali merasa perlu mengikuti pelajaran tambahan dari lembaga bimbingan belajar di luar sekolah (bimbel) yang menekankan pada latihan soal dan pencapaian dalam ujian tertulis (Subkhan, 2019). Hal ini mempersempit ruang bagi penerapan metode pengajaran yang lebih bermakna, kontekstual, dan kreatif seperti pembelajaran berbasis problem atau *project*. Fenomena inilah yang mendorong negara-negara Asia Timur dengan sistem pendidikan yang baik (seperti Hong Kong, Taiwan, Korea, dan Singapura) untuk mengurangi penggunaan tes terstandar dan mendorong guru untuk menerapkan asesmen yang lebih inovatif dan holistik (Zhao, 2015).

## 5.3 Demoralisasi dan dampak psikologis lain

Sebagai pekerja profesional, guru seharusnya memiliki agensi yang kuat untuk menjalankan tugasnya secara otonom. Di Indonesia, survei internal yang dilakukan PASKA Kemendikbud menunjukkan bahwa banyak guru yang merasa tidak berdaya untuk memotivasi siswa jika UN dan USBN dihilangkan (PASKA Kemendikbud, 2019). Survei yang sama juga menemukan bahwa banyak guru yang bingung dengan cara menilai hasil belajar siswa. Tujuan pengajaran tereduksi menjadi keberhasilan mendapat skor tinggi pada UN dan USBN. Fenomena ini juga terjadi di berbagai negara lain yang menerapkan tes-tes terstandar. Akuntabilitas berdasarkan tes-tes terstandar dipersepsi secara luas oleh guru sebagai tekanan yang menciptakan budaya ketakutan (Emler et al., 2019). Kegagalan untuk mendapat nilai baik pada tes-tes tersebut dapat menciderai kesejahteraan psikologis dan kepuasan kerja guru (Abu-alhija, 2007; Berliner, 2011; Emler et al., 2019; Erskine, 2014; Klenowski, 2011). Tekanan tes-tes terstandar juga berdampak negatif pada kesejahteraan psikologis dan motivasi intrinsik siswa untuk belajar (Gregory & Clarke, 2010; Wyn et al., 2014; Zhao, 2015).

## 5.4 Memicu praktik curang

Risiko gagal dalam ujian yang berisiko tinggi dapat mendorong praktik curang, terutama pada mereka yang memiliki motivasi belajar dan efikasi (keyakinan diri) rendah untuk mencapai hasil yang memuaskan dalam ujian (Finn & Frone, 2004). Kecurangan dalam ujian-ujian *high stakes* di Cina, seperti seleksi masuk perguruan tinggi dan ujian menjadi pegawai pemerintah, sudah sering dilaporkan (Kirkpatrick & Zang, 2011; Moore, 2013; Yu & Suen, 2005). Di Amerika, tekanan tes-tes terstandar juga mendorong kecurangan secara sistematis oleh guru dan administrator sekolah. Kasus yang paling terkenal yaitu yang terjadi di distrik Atlanta yang terkuak pada 2008. Ketika itu, ratusan guru mengaku berkoordinasi untuk meningkatkan skor dengan cara-cara seperti mengkoreksi jawaban siswa. Kecurangan tersebut telah terjadi sejak bertahun-tahun sebelumnya (Osunsami & Forer, 2011). Di Indonesia, kecurangan terkait UN juga terjadi dalam bentuk penyebaran paket-paket soal dan kunci jawabannya sebelum ujian dilangsungkan (Sujatmiko, 2014; Wurinanda, 2016). Kemendikbud berupaya mengurangi kecurangan dengan cara menambah jumlah paket soal dan menerapkan ujian berbasis komputer (Akuntono, 2012; Batubara, 2016). Upaya-upaya tersebut tampaknya berhasil mengurangi, namun tidak sepenuhnya menghilangkan kecurangan (Abdi, 2019).

## 5.5 Melanggengkan kesenjangan

Akibat ketimpangan struktur sosial-ekonomi di masyarakat, siswa dari keluarga miskin cenderung memiliki keterampilan akademik yang lebih buruk dan kurang siap untuk belajar secara akademik (Davies & Rizk, 2017; Roksa & Potter, 2011; Lee & Burkam, 2002). Dengan kata lain, siswa dari keluarga miskin cenderung mendapatkan kesempatan belajar yang lebih sedikit, dan lebih buruk, dibanding rekan mereka dari keluarga menengah-atas (OECD, 2018; Schmidt, Burroughs, Zoido, & Houang, 2015). Di Indonesia, kesenjangan kesempatan belajar ini terjadi di sekolah maupun di luar sekolah, di mana akses pada bimbingan belajar tergantung pada sumber daya finansial yang dimiliki keluarga siswa. Kesenjangan kesempatan belajar ini tercermin juga pada performa siswa dalam tes-tes terstandar, di mana siswa dari keluarga miskin cenderung mendapat skor lebih rendah (Zwick & Green, 2007). Karena itu, penggunaan tes-tes terstandar untuk menentukan kelulusan dan seleksi berpotensi melanggengkan kesenjangan (Emler et al., 2019).

Dengan demikian, asesmen nasional seharusnya dirancang bukan hanya untuk menilai pencapaian standar nasional, tapi juga mendorong terciptanya praktik asesmen kelas yang berorientasi pada kemajuan belajar siswa. Rancangan asesmen nasional juga perlu dilandasi kesadaran akan dampak negatif yang tidak diinginkan seperti disebutkan sebelumnya.

# 6. Reformasi Sistem Asesmen: Rekomendasi Kebijakan

## 6.1 Prinsip Reformasi Sistem Asesmen

Asesmen nasional untuk mengevaluasi sistem pada akhirnya bertujuan membawa perubahan positif dalam proses belajar-mengajar yang dipraktikkan di kelas. Namun asesmen nasional tersebut juga bisa memiliki efek samping yang negatif, seperti diuraikan pada bagian sebelumnya (termasuk penyempitan kurikulum, distorsi metode pengajaran, demoralisasi, dan tekanan untuk curang). Berikut ialah beberapa prinsip yang berguna untuk meminimalkan kemungkinan terjadinya efek samping tersebut:

- ▶ **Asesmen untuk evaluasi sistem seyogyanya dipisahkan dari evaluasi siswa (termasuk kelulusan dan seleksi).** Dengan demikian, asesmen nasional untuk evaluasi sistem dapat dirancang agar tidak akan memiliki konsekuensi apa pun bagi siswa, dan keikutsertaan dalam asesmen tidak menjadi beban psikologis bagi mereka. Pemisahan ini juga penting agar asesmen siswa tidak terikat oleh keterbatasan praktis yang melekat pada asesmen skala besar (seperti bentuk soal penilaian tertulis yang kemudian secara praktis didominasi oleh pilihan ganda dan isian singkat).
- ▶ **Asesmen untuk evaluasi sistem sebaiknya dilakukan pada sampel dan tidak setiap tahun.** Asesmen pada sampel siswa menegaskan bahwa hasilnya tidak mungkin memiliki konsekuensi pada level individu. Selain itu, asesmen juga bisa dilakukan setiap dua tahun sekali, mengingat bahwa perubahan kebijakan pendidikan membutuhkan waktu cukup panjang sebelum dampaknya bisa terlihat. Pelaksanaan asesmen setiap dua tahun ini akan mengurangi tekanan kepada guru dan sekolah, sehingga mengurangi keinginan atau kebutuhan untuk berlaku curang.
- ▶ **Pelaporan hasil menekankan pada informasi untuk perbaikan dan peningkatan mutu, bukan komparasi antarsekolah dan daerah.** Setiap siswa memiliki latar belakang sosial ekonomi dan sumber daya pendukung di lingkungan keluarga yang berbeda. Sementara itu, prestasi siswa dalam asesmen terstandar dipengaruhi oleh faktor-faktor tersebut – yang sayangnya berada di luar kendali sekolah. Laporan hasil hendaknya merefleksikan mutu proses belajar-mengajar di sekolah. Karena itu, laporan hasil asesmen menekankan pada perubahan capaian di masing-masing sekolah. Hal ini akan lebih mendorong guru dan sekolah untuk fokus pada upaya perbaikan internal dan berkolaborasi, bukan bersaing dengan sekolah-sekolah lain yang memiliki kondisi dan sumber daya berbeda.

- ▶ **Pelaporan hasil tidak semata menunjukkan capaian belajar siswa namun disertai dengan penjelasan faktor input maupun proses yang memengaruhi.** Hasil belajar dipengaruhi oleh banyak aspek, baik yang dapat diintervensi oleh sekolah maupun yang di luar kendali sekolah. Maka evaluasi sistem pendidikan yang baik mampu memberikan informasi mengenai faktor-faktor capaian hasil belajar yang dapat diupayakan sekolah. Faktor-faktor tersebut menggambarkan kualitas pembelajaran di sekolah, iklim sekolah yang kondusif untuk siswa belajar, tingkat kepedulian dan keterlibatan warga sekolah dalam memajukan sekolah. Sistem evaluasi yang memberikan informasi berupa input, proses maupun hasil dapat memberikan kesempatan kepada setiap kepala sekolah untuk melakukan perencanaan perbaikan yang berbasis pada data.

Selain meminimalkan risiko munculnya efek samping negatif, asesmen nasional untuk evaluasi sistem juga perlu dirancang untuk memaksimalkan potensi dampak positifnya bagi proses belajar mengajar. Berikut beberapa prinsip yang bermanfaat untuk itu:

- ▶ **Fokus asesmen hendaknya pada pengukuran kemampuan atau kapabilitas mendasar yang bisa diterapkan secara luas di berbagai situasi.** Kapabilitas semacam ini beririsan dengan apa yang kerap disebut sebagai “kompetensi abad 21” (Griffin, McGaw, & Care, 2012). Contohnya yaitu kemampuan bernalar tentang teks (literasi) dan angka (numerasi), disposisi untuk belajar sepanjang hayat, keterampilan berkomunikasi dan bekerja sama dalam konteks multikultural, serta kecakapan berpikir kritis secara etis (Gardner, 2008).

Isu relevansi kemampuan atau kapabilitas tersebut menjadi sangat penting. Saat ini seringkali yang menjadi target asesmen ialah penguasaan siswa akan konten atau materi kurikulum. Hal ini menjadi problematik karena tidak seluruh konten kurikulum relevan bagi siswa. Di era informasi saat ini, pengetahuan faktual semakin mudah diperoleh dan diakses oleh hampir setiap orang, sehingga sekedar mengetahui menjadi tidak cukup dan kurang relevan. “...its not longer about what you know, but what you can do with what you know...” (Schleicher, 2018). Terlebih lagi, mengukur pengetahuan tentang konten dengan cakupan yang luas cenderung mendorong pengajaran yang berorientasi pada “penyelesaian materi” secara dangkal.

Fokus pengukuran saat ini perlu bergeser pada kemampuan siswa untuk menggunakan dan mengevaluasi pengetahuan untuk merumuskan serta menyelesaikan masalah, baik yang bersifat rutin maupun non-rutin (belum pernah atau jarang ditemui). Dengan menggeser fokus dari keluasan pengetahuan menuju kedalaman kompetensi, asesmen nasional bisa mendorong guru untuk lebih banyak menerapkan metode-metode pengajaran yang membuat siswa berpikir kritis dan kreatif.

- ▶ **Asesmen hendaknya memberi informasi tentang posisi siswa dalam trajektori atau lintasan belajar berjangka panjang.** Kemampuan atau kompetensi yang berharga untuk dipelajari hampir selalu memerlukan waktu panjang untuk dikuasai. Ini berlaku juga untuk kompetensi seperti literasi membaca, menulis, numerasi, kemampuan bekerja sama, dst. Kompetensi membaca, misalnya, mencakup kemampuan dasar seperti mengenali huruf, sampai kemampuan mengevaluasi dan menyintesis informasi dari berbagai sumber yang mungkin saling bertentangan (OECD, 2016). Asesmen seharusnya menjadi alat untuk membantu siswa (dan guru) melihat di mana ia berada dalam perjalanannya menjadi seseorang ahli di sebuah area kompetensi. Asesmen seharusnya membantu siswa dan guru memahami apa yang bisa dilakukan seorang yang ahli. Asesmen juga semestinya membantu siswa dan guru untuk memantau kemajuan belajar dalam perjalanannya menjadi lebih ahli di area kompetensi tersebut.
- ▶ **Guru dan sekolah perlu dibantu agar dapat memanfaatkan hasil asesmen.** Asesmen perlu dirancang agar hasilnya dapat berfungsi sebagai umpan balik bagi perbaikan mutu belajar mengajar. Untuk itu, hasil asesmen perlu disampaikan dalam waktu yang cukup memadai agar masih dipandang relevan dan dapat dimanfaatkan oleh pengguna. Saat ini, akibat pencampuradukkan fungsi penilaian, hasil UN sebagai asesmen nasional sulit dimanfaatkan guru dan sekolah untuk perbaikan di level siswa karena diujikan di akhir jenjang sekolah dan siswa yang menerima hasil telah lulus. Meskipun hasil ujian nasional telah diolah sampai analisis di level cakupan materi bahkan butir soal, hal ini akan sulit dimanfaatkan jika guru dan sekolah terbatas kapasitasnya untuk memahami, memaknai, dan menindaklanjuti hasil asesmen.
- ▶ **Asesmen adalah bagian yang tidak lepas dari standar (*intended curriculum*) serta pembelajaran (*implemented curriculum*).** Asesmen memonitor seberapa pembelajaran berhasil mengantarkan siswa mencapai standar yang telah ditetapkan, serta sekaligus mengevaluasi tingkat ketercapaian (*achievable*) dan tingkat kepraktisan (*deliverable*) standar yang ditetapkan. Idealnya, desain asesmen serta hasil asesmen bersifat selaras (*aligned*) dengan standar/kurikulum dan praktik pembelajaran. Meski demikian, rekomendasi reformasi sistem asesmen ini tidak secara mutlak dibatasi oleh kurikulum dan praktik pengajaran yang saat ini berlaku. Keselarasan antara asesmen dan kurikulum menuntut reformasi pada keduanya secara bersamaan.

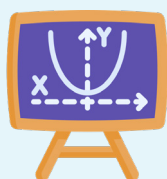
Prinsip-prinsip ini perlu diterjemahkan menjadi kebijakan dan sistem yang spesifik. Bagian berikut merupakan rekomendasi mengenai salah satu alternatif penerjemahan prinsip-prinsip tersebut.



## 6.2 Rekomendasi Sistem Asesmen Nasional untuk Evaluasi Sistem

Sebuah sistem asesmen dengan tiga komponen direkomendasikan sebagai sebuah alternatif pengganti UN (lihat Tabel 3). Komponen asesmen alternatif pengganti UN pertama dan kedua disebut sebagai “Asesmen Kompetensi Minimum” (AKM) karena mengukur kompetensi inti yang diharapkan bisa dikuasai oleh semua atau hampir semua (*vast majority*) siswa di Indonesia. Dalam hal ini, minimum tidak berarti standar yang rendah. Komponen ketiga (Survei Karakter dan Survei Lingkungan Belajar) mengukur luaran belajar yang lebih bersifat sosial-emosional, serta kualitas proses belajar-mengajar di tiap sekolah.

Sistem asesmen yang diusulkan ini memiliki beberapa karakteristik yang dirancang agar berdampak positif pada proses belajar-mengajar dan lebih berpihak pada siswa.



**Pertama**, AKM akan mengukur literasi membaca dan literasi matematika (numerasi). Kedua literasi ini merupakan kecakapan dasar terpenting untuk bisa belajar sepanjang hayat dan berkontribusi secara produktif dalam kehidupan modern. Yang dimaksud literasi di sini adalah kompetensi berpikir atau bernalar siswa ketika membaca teks dan menghadapi persoalan yang membutuhkan pengetahuan matematika. AKM tidak dirancang untuk mengukur pengetahuan tentang kurikulum secara detail dan rinci. AKM diharapkan menggeser fokus tujuan belajar dari pengetahuan faktual secara luas ke arah penguasaan kompetensi mendasar yang bermanfaat secara luas.

Konstrak yang diukur pada AKM dimulai dari dua literasi fundamental namun seiring waktu dapat diperluas untuk mencakup literasi dasar lainnya, seperti literasi menulis, literasi sains, literasi sosial budaya, literasi digital, serta literasi finansial. Hal ini mempertimbangan aspek teknis pelaksanaan asesmen skala nasional yang memerlukan alat ukur terstandar yang kredibel serta akuntabel. Penambahan konstrak literasi lainnya juga akan dilakukan setelah praktik baik pemanfaatan hasil AKM literasi membaca dan literasi matematika terpola dan tersistem secara efektif dan efisien. Karena poin terpenting dari asesmen ialah pemanfaatan hasil dari informasi yang minimal sekalipun, bukan mengenai seberapa banyak dan luas konstrak asesmen yang dilakukan.

Tabel 3. Komponen dan karakteristik asesmen pengganti UN

	Asesmen Kompetensi Minimum – versi kelas	Asesmen Kompetensi Minimum - versi nasional	Survei Karakter dan Survei Lingkungan Belajar
<i>Fungsi</i> <i>Yang diukur</i>	Memahami hasil belajar individu siswa 1. Literasi membaca 2. Literasi matematika	Mengevaluasi kualitas sistem pendidikan 1. Literasi membaca 2. Literasi matematika	Mengevaluasi kualitas sistem pendidikan 1. Disposisi dan keterampilan sosial-emosional-etis-spiritual siswa 2. Kesejahteraan psikologis siswa dan guru 3. Praktik-praktik pengajaran/iklim belajar, iklim sekolah, partisipasi dan keterlibatan warga sekolah
<i>Peserta</i>	Siswa kelas 2 sampai 12	Sampel dari: <ul style="list-style-type: none"><li>siswa kelas 4 atau 5 SD/MI sederajat</li><li>siswa kelas 8 SMP/MTs sederajat</li><li>siswa kelas 11 SMA/MA dan SMK sederajat</li></ul>	Sampel dari: <ul style="list-style-type: none"><li>siswa kelas 4 atau 5 SD/MI sederajat</li><li>siswa kelas 8 SMP/MTs sederajat</li><li>siswa kelas 11 SMA/MA dan SMK sederajat</li><li>Kepala sekolah</li><li>guru</li></ul>
<i>Pelaksanaan</i>	Sepenuhnya mandiri oleh guru sebagai bagian dari proses belajar-mengajar	Terpusat di bawah koordinasi Kemendikbud; situasi asesmen terstandar	Terpusat di bawah koordinasi Kemendikbud; situasi asesmen terstandar. Di tingkat satuan pendidikan, disediakan instrumen yang mampu digunakan secara mandiri oleh guru (dan/atau satuan pendidikan) untuk melakukan observasi dalam rangka menilai seberapa membudaya karakter yang diharapkan.
<i>Frame of reference untuk memaknai hasil</i>	1. Model konseptual yang menggambarkan tingkatan kecakapan ( <i>proficiency</i> ) dalam literasi membaca dan literasi matematika. Siswa mengetahui posisi kompetensi dirinya dalam suatu skala belajar kontinum.	1. Model konseptual yang menggambarkan tingkatan kecakapan ( <i>proficiency</i> ) dalam literasi membaca dan literasi matematika. 2. Skor sekolah/daerah yang sama di masa lampau.	1. Skor sekolah/daerah yang sama di masa lampau. 2. Standar minimum yang bisa ditetapkan berdasarkan pertimbangan normatif dan praktis sesuai kondisi setiap sekolah atau daerah. 3. Profil perkembangan karakter individu siswa yang diperoleh dari penilaian guru melalui observasi berulang. 4. Profil iklim sekolah berdasarkan kategori yang ditetapkan berdasarkan pertimbangan normatif dan reflektif.

	Asesmen Kompetensi Minimum – versi kelas	Asesmen Kompetensi Minimum - versi nasional	Survei Karakter dan Survei Lingkungan Belajar
	2. Skor individu siswa yang sama di masa lampau.	3. Standar minimum yang bisa ditetapkan berdasarkan pertimbangan normatif dan praktis sesuai kondisi setiap sekolah atau daerah.	
<i>Penggunaan hasil</i>	Memantau dan mendorong kemajuan belajar individu siswa	Alokasi fasilitas, sumber daya manusia, dan pelatihan sesuai kebutuhan sekolah atau daerah	Alokasi fasilitas, sumber daya manusia, dan pelatihan sesuai kebutuhan sekolah atau daerah
<i>Pelaporan hasil</i>	Hasil level individu siswa dapat dilaporkan secara informal kepada siswa ybs, orangtua, kolega guru, dan/ atau kepala sekolah sesuai keperluan	<ol style="list-style-type: none"> <li>1. Hasil level sekolah dilaporkan kepada sekolah ybs dan dinas setempat (tidak secara publik)</li> <li>2. Hasil level dinas/daerah dilaporkan secara publik dengan menekankan konteks sosial-ekonomi dan sumber daya, serta <i>progress</i> antarwaktu (bukan perbandingan/ peringkat antardaerah)</li> </ol>	<ol style="list-style-type: none"> <li>1. Hasil level sekolah dilaporkan kepada sekolah ybs dan dinas setempat (tidak secara publik)</li> <li>2. Hasil level dinas/daerah dilaporkan secara publik dengan menekankan konteks sosial-ekonomi dan sumber daya, serta <i>progress</i> antar waktu (bukan perbandingan/peringkat antardaerah)</li> <li>3. Hasil observasi siswa (oleh guru) dilaporkan kepada siswa ybs dan orangtua sebagai komitmen bersama antara guru dan lingkungan keluarga dalam menumbuhkan karakter siswa</li> </ol>



**Kedua**, sistem yang diusulkan juga mengukur dimensi-dimensi sosial-emosional mencirikan karakter siswa. Hal ini penting untuk menyampaikan pesan bahwa proses belajar-mengajar harus mengembangkan potensi siswa secara utuh. Karakteristik sosial-emosional apa saja yang perlu diukur memerlukan kajian akademik tersendiri.

**Ketiga**, sistem ini mencakup Survei Lingkungan Belajar untuk mendapatkan informasi tentang kualitas pengajaran dan iklim sekolah. Ini mencakup variabel-variabel seperti metode pengajaran yang biasa diterapkan guru, pemenuhan kebutuhan-kebutuhan psikologis dasar siswa dan guru, dan kualitas interaksi antarsiswa dan antara siswa dengan guru di sekolah (termasuk insiden *bullying*). Informasi ini diperlukan untuk memberi petunjuk pedagogis yang relevan bagi perbaikan mutu belajar-mengajar di masing-masing sekolah.



**Keempat**, sistem ini mencakup AKM dalam versi yang dapat digunakan oleh guru secara mandiri. AKM Kelas ini khusus dirancang untuk digunakan guru dalam proses belajar-mengajarnya secara mandiri. AKM Kelas sebaiknya dibuat untuk semua jenjang kelas, mulai dari SD awal (setidaknya kelas 2) sampai akhir SMA. Jika AKM nasional tidak melaporkan kemampuan individu siswa, maka AKM kelas merupakan alat diagnosa di level individu yang menjadi kewenangan guru. Penggunaan AKM Kelas diharapkan menumbuhkan keyakinan diri pada guru untuk membantu siswanya menghadapi asesmen versi nasional. AKM Kelas tersebut juga diharapkan menjadi inspirasi bagi guru-guru untuk mengembangkan penilaian formatif yang berkualitas secara mandiri sesuai dengan kebutuhan guru secara spesifik.

**Kelima**, instrumen-instrumen asesmen harus dilandasi oleh model *learning progression* yang menggambarkan tahap-tahap perubahan dari *novice* (pemula) menjadi *expert* (ahli) dalam area kompetensi yang relevan. Dengan demikian, asesmen akan membantu siswa dan guru melihat posisinya dalam perjalanan belajar berjangka panjang, bukan untuk menghakimi keberhasilan atau kegagalan dalam menguasai penguasaan materi atau konten yang dianggap standar pada jenjang tertentu. Hal ini penting karena setiap siswa, meski berada di kelas yang sama, memiliki titik start dan kecepatan belajar yang berbeda. Pemaknaan hasil asesmen sebagai keberhasilan/kegagalan mencapai standar yang sama bagi tiap siswa mengabaikan fakta ini.

**Keenam**, peserta asesmen ialah siswa kelas 4/5, 8, dan 11. Hasil asesmen diharapkan bisa diberikan kepada guru/sekolah ketika peserta asesmen masih menjadi siswa di sekolah tersebut. Dengan demikian, hasil asesmen dapat dimanfaatkan guru untuk mengembangkan pengajaran sesuai kebutuhan belajar siswa di kelasnya. Untuk SMP dan SMA, rekomendasi pemilihan kelas 8 dan 11 dilandasi pemikiran bahwa salah satu tujuan asesmen yaitu untuk mengukur efektivitas proses belajar-mengajar di tiap sekolah. Hal ini tidak bisa dicapai jika asesmen dilakukan pada siswa kelas 7 dan 10 yang baru saja mengikuti pembelajaran di SMP dan SMA.

**Ketujuh**, pelaporan hasil AKM (versi nasional), Survei Karakter, dan Survei Lingkungan Belajar disampaikan kepada masing-masing sekolah dan dinas pendidikan. Hasil asesmen dilaporkan pada level agregat (sekolah dan dinas), bukan skor untuk masing-masing siswa. Laporan asesmen juga sebaiknya menekankan pada *progress* atau perubahan kinerja sekolah, bukan perbandingan dan persaingan antarsekolah/dinas. Karena itu hasil asesmen ini tidak perlu dipublikasikan secara terbuka untuk menghindari potensi penggunaan yang kontra-produktif seperti perangkingan sekolah dan dinas oleh media atau pihak lain. Hasil tahun pertama dimaknai semata-mata sebagai *baseline*, bukan pencapaian atau prestasi.

## 6.3 Rekomendasi mengenai asesmen untuk kelulusan dan seleksi

Dengan dipisahkannya asesmen untuk evaluasi sistem dan evaluasi siswa, fungsi evaluasi siswa yang tadinya dilayani oleh UN dan USBN (Ujian Sekolah Berstandar Nasional) beralih menjadi tanggung jawab guru dan sekolah. Dua fungsi yang krusial asesmen untuk evaluasi pada level siswa yaitu (1) penentuan kelulusan (pemberian ijazah sebagai pengakuan penyelesaian sebuah jenjang sekolah) dan (2) seleksi masuk jenjang sekolah yang lebih tinggi.

Pada prinsipnya, yang mengetahui dan berwenang menilai hasil belajar siswa ialah guru. Hasil belajar selama 6 tahun (untuk SD) atau 3 tahun (untuk SMP dan SMA) terlalu luas dan kompleks untuk bisa diukur hanya menggunakan tes tertulis yang diselenggarakan terpusat. Hanya guru – melalui asesmen yang berkelanjutan menggunakan beragam instrumen – yang bisa melakukan penilaian pencapaian belajar siswa secara komprehensif. Sejalan dengan prinsip ini, UU Sisdiknas menetapkan penilaian siswa sebagai kewenangan guru (Pasal 58 ayat 1), dan penentuan kelulusan sebagai kewenangan sekolah (Pasal 61). Meski dilakukan oleh guru dan diselenggarakan oleh masing-masing sekolah, asesmen untuk kelulusan tentu perlu mengacu pada tujuan belajar yang dirumuskan dalam kurikulum sekolah maupun Standar Nasional Pendidikan.

Tabel 4. Perbandingan karakteristik asesmen untuk kelulusan dan untuk seleksi

	Kelulusan	Seleksi
Yang berwenang	Sekolah asal	Sekolah/institusi sasaran, dinas pendidikan, dan pemerintah pusat
Cakupan asesmen	Seluruh kompetensi yang ditetapkan sebagai tujuan belajar dalam kurikulum yang digunakan sekolah asal	1. Kompetensi yang dinilai berharga oleh sekolah/institusi sasaran, dan/atau 2. Kompetensi yang memprediksi prestasi masa depan di sekolah/institusi sasaran

	Kelulusan	Seleksi
<i>Landasan interpretasi hasil asesmen</i>	Standar pencapaian ( <i>proficiency</i> ) yang ditetapkan dalam kurikulum sekolah asal	1. Perbandingan normatif dengan populasi skor peserta seleksi 2. Kerangka asesmen yang melandasi masing-masing tes yang digunakan untuk seleksi
<i>Komparabilitas hasil asesmen</i>	Keterbandingan kualitatif (kategori pencapaian) antarsiswa dalam satu sekolah	Keterbandingan kuantitatif antarsiswa yang mendapat di sekolah/institusi sasaran
<i>Peserta asesmen</i>	Semua siswa yang berada di akhir masa studi	Siswa yang berminat mengikuti seleksi
<i>Instrumen yang sesuai</i>	Kombinasi antara beragam instrumen (terstandar maupun tidak), sesuai dengan sifat kompetensi yang diukur dan kebutuhan/profil belajar siswa	1. Instrumen terstandar (misalnya, ujian tertulis dan asesmen kinerja dengan rubrik penskoran yang terstandar) 2. Sekolah/institusi dengan selektivitas tinggi memerlukan instrumen yang sensitif untuk membedakan siswa pada <i>range</i> kemampuan yang tinggi

Seleksi memasuki jenjang sekolah yang lebih tinggi diatur dalam regulasi penerimaan siswa baru. Karena negara berkomitmen menjamin hak belajar siswa sampai 9 tahun, pendaftaran SD dan SMP seharusnya didasarkan pada mekanisme zonasi dan afirmasi, bukan seleksi prestasi. Daerah-daerah yang belum memiliki jumlah bangku SMP yang memadai didorong untuk menambah sekolah atau rombongan belajar di SMP yang ada. Namun sebelum kondisi tersebut tercapai, akan ada kebutuhan untuk melakukan seleksi calon siswa SMP berdasarkan prestasi. Hal ini berlaku juga untuk seleksi calon siswa SMA.

Idealnya, seleksi dilakukan berdasarkan hasil asesmen di sekolah asal (nilai rapor). Hal ini dapat mengurangi beban biaya dan waktu bagi siswa/keluarga maupun sistem. Namun hal ini tidak selalu dapat dilakukan karena asesmen untuk kelulusan dan seleksi memiliki karakteristik yang berbeda (*lihat Tabel 4*). Hambatan utama penggunaan asesmen kelulusan untuk seleksi ialah soal *comparability*: nilai rapor dari sekolah-sekolah asal sulit digunakan sebagai landasan memeringkat prestasi siswa secara adil. Untuk mengakomodir sistem seleksi yang berkeadilan, transparan, serta akuntabel, pemerintah hendaknya mendorong kapasitas dinas/wilayah dalam melakukan penilaian terstandar. Kapasitas dinas ini penting sehingga mampu menyelenggarakan sistem seleksi yang kredibel dan akuntabel, tanpa mencampuradukkan fungsinya dengan asesmen kelulusan. Kesadaran dinas untuk memisahkan sistem seleksi dari sistem kelulusan perlu ditingkatkan, tanpa mengabaikan tujuan utama pemerataan mutu semua satuan pendidikan.

Dalam jangka panjang, seiring dengan meningkatnya kapasitas profesional guru, semakin banyak guru yang mampu menerjemahkan standar nasional dengan baik dan mengukurnya secara valid. Nilai rapor dari sekolah yang berbeda-beda akan menjadi lebih bisa dibandingkan. Untuk itu, pengembangan kompetensi guru untuk menerapkan asesmen perlu menjadi bagian dari kebijakan Kemendikbud mengenai pendidikan guru. Selain itu, Kemendikbud juga perlu memikirkan pengembangan beragam instrumen dan prosedur asesmen yang terstandar dan bermutu tinggi yang dapat digunakan secara mandiri oleh guru untuk menentukan kelulusan siswanya.

## 7. Penutup

Di banyak negara, asesmen kerap digunakan sebagai instrumen untuk menggerakkan perubahan dalam aspek-aspek lain dari sebuah sistem pendidikan. Namun pengalaman berbagai negara, termasuk juga Indonesia, menunjukkan bahwa pengaruh asesmen tidak selalu positif. Berbagai penelitian bahkan mengindikasikan bahwa ketika hasilnya digunakan sebagai penentu keberhasilan siswa, guru, dan/atau sekolah, asesmen berpotensi menjadi tekanan yang menyempitkan kurikulum dan membuat pengajaran menjadi dangkal. Dalam lingkungan belajar seperti itu, siswa kerap menjadi pihak yang paling dirugikan.

Rekomendasi yang disampaikan dalam laporan ini merupakan kerangka kebijakan asesmen yang lebih berpihak pada siswa. Dalam kerangka kebijakan ini, asesmen dirancang untuk melayani fungsi formatif pada level sistem. Dengan kata lain, asesmen dimaksudkan sebagai alat untuk menghasilkan informasi yang bermanfaat bagi perbaikan sistem, mulai dari level kelas, sekolah, daerah, dan nasional. Asesmen hendaknya memberikan tantangan yang “pas dosisnya kepada siswa”. Asesmen memberikan tingkat tantangan yang tepat untuk siswa selalu termotivasi belajar, namun tidak memberikan tekanan psikologis. Dan sebaliknya, asesmen juga tidak sedemikian mudah sehingga membosankan dan menenggelamkan daya juang siswa. Dampak akhir yang diharapkan ialah meningkatnya mutu proses belajar-mengajar bagi semua siswa. Guru menjadi fasilitator yang bijak bagi setiap siswa dalam mengarungi perjalanan belajarnya masing-masing.

Perlu digarisbawahi bahwa dampak asesmen akan bergantung pada bagaimana hasilnya digunakan oleh para pengambil kebijakan serta para guru di sekolah. Asesmen yang dirancang untuk bersifat *low-stakes* dan formatif pun dapat digunakan semata-mata sebagai cambuk persaingan antarsekolah atau sekadar untuk melabeli sekolah sebagai “gagal” atau “berhasil”. Untuk itu, penyesuaian paradigma perlu dilakukan agar guru, sekolah, dan dinas pendidikan memaknai reformasi asesmen sebagai upaya untuk menjadi lebih berpihak pada siswa.

## 8. Pustaka Acuan

- Abdi, A. P. (2019, May 7). Kemendikbud Catat 126 Kecurangan Selama Ujian Nasional 2019. *Tirto.id*. Retrieved from <https://tirto.id/kemendikbud-catat-126-kecurangan-selama-ujian-nasional-2019-drNd>.
- Abu-alhija, F. N. (2007). Large-scale testing: benefits and pitfalls. *Studies in Educational Evaluation*, 33, 50–68. <https://doi.org/10.1016/j.stueduc.2007.01.005>.
- Akuntono, I. (2012, September 16). Tekan Kecurangan, 20 Paket Soal untuk UN 2013. *Kompas.Com*. Retrieved from <https://edukasi.kompas.com/read/2012/09/16/11385057/tekan-kecurangan.20.paket.soal.untuk.un.2013>.
- Ashadi, A., & Rice, S. (2016). High stakes testing and teacher access to professional opportunities: lessons from Indonesia. *Journal of Education Policy*, 31(6), 727–741. <https://doi.org/10.1080/02680939.2016.1193901>.
- Batubara, H. (2016, April 4). Mendikbud: Pelaksanaan UN Berbasis Komputer Meningkatkan Drastis Dalam Setahun. *Detik News*. Retrieved from <https://news.detik.com/berita/d-3178707/mendikbud-pelaksanaan-un-berbasis-komputer-meningkat-drastis-dalam-setahun>.
- Berliner, D. (2011). Rational responses to high stakes testing: the case of curriculum narrowing and the harm that follows. *Cambridge Journal of Education*, 41(3), 37–41. <https://doi.org/10.1080/0305764X.2011.607151>.
- Berry, R. (2011). Educational Assessment in Mainland China, Hong Kong and Taiwan. In R. Berry & B. Adamson (Eds.), *Assessment Reform in Education: Policy and Practice* (pp. 49–62). Dordrecht: Springer Science+Business Media.
- Biggs, J. (1996). *Testing: To educate or to select? Hong Kong: Hong Kong Educational Publishing Co.* Hong Kong: Hong Kong Educational Publishing Co.
- Carless, D., & Carless, D. (2005). Prospects for the implementation of assessment for learning. *Assessment in Education: Principles, Policy & Practice*, 12(1), 39–54. <https://doi.org/10.1080/0969594042000333904>.
- Curriculum Development Council (CDC). (2002). *Basic education curriculum guide: Building on strengths (Primary 1 – Secondary 3)*. Hong Kong: Curriculum Development Council (CDC).
- Davies, S., & Rizk, J. (2017). The Three Generations of Cultural Capital Research: A Narrative Review. *Review of Educational Research*, 88(3), 003465431774842. <https://doi.org/10.3102/0034654317748423>.
- Dee, T. S., & Jacob, B. (2011). The Impact of No Child Left Behind on Student Achievement. *Journal of Policy Analysis and Management*, 30(3), 418–446. <https://doi.org/10.1002/pam>.
- Emler, T. E., Zhao, Y., & Wang, Y. (2019). Side Effects of Large-Scale Assessments in Education. *ECNU Review of Education*, 2(3), 279–296. <https://doi.org/10.1177/2096531119878964>.
- Erskine, J. L. (2014). It Changes How Teachers Teach: How Testing Is Corrupting Our Classrooms and Student Learning. *Multicultural Education*, 21(2), 38–40.
- Finn, K. V., & Frone, M. R. (2004). Academic Performance and Cheating: Moderating Role of School Identification and Self-Efficacy. *The Journal of Educational Research*, 87(3), 115–121. <https://doi.org/10.3200/JOER.97.3.115-121>.
- Gable, A., & Lingard, B. (2013). *NAPLAN and the Performance Regime in Australian Schooling: A review of the Policy Context* (Research Paper). Brisbane, Australia.



- Gardner, H. (2008). *Five Minds for the Future*. Cambridge, MA: Harvard Business School Press.
- Gopinathan, S. (1999). Preparing for the Next Rung: economic restructuring and educational reform in Singapore. *Journal of Education and Work*, 12(3), 295–308. <https://doi.org/10.1080/1363908990120305>.
- Goss, P. (2019, December 10). What next for Australia's education system after abysmal PISA results? *Grattan Institute*. Retrieved from <https://grattan.edu.au/news/what-next-for-australias-education-system-after-abysmal-pisa-results/>.
- Gregory, K., & Clarke, M. (2010). High-Stakes Assessment in England and Singapore. *Theory Into Practice*, 42(1), 66–74. <https://doi.org/10.1207/s15430421tip4201>.
- Griffin, P., McGaw, B., & Care, E. (2012). *Assessment and Teaching of 21st Century Skills*. Dordrecht: Springer.
- Hanushek, E. A., & Raymond, M. E. (2005). *Does School Accountability Lead to Improved Student Performance* (NBER Working Paper Series). Massachusetts.
- Hoffmann, J. (2003). *History of the No Child Left Behind Act of 2001 (NCLB)* (Vol. Assessment).
- Howell, A. (2017). 'Because then you could never ever get a job!': children's constructions of NAPLAN as high-stakes. *Journal of Education Policy*, 32(5), 564–587. <https://doi.org/10.1080/02680939.2017.1305451>.
- Huntington-klein, N., & Ackert, E. (2018). The Long Road to Equality: A Meta-Regression Analysis of Changes in the Black Test Score Gap Over Time. *Social Science Quarterly*. <https://doi.org/10.1111/ssqu.12483>.
- Kennedy, K. J., Chan, J. K. S., Fok, P. K., & Yu, M. W. (2008). Forms of assessment and their potential for enhancing learning: conceptual and cultural issues. *Educational Research: Policy and Practice*, 7, 197–207. <https://doi.org/10.1007/s10671-008-9052-3>.
- Kirkpatrick, R., & Zang, Y. (2011). The Negative Influences of Exam-Oriented Education on Chinese High School Students: Backwash from Classroom to Child. *Language Testing in Asia*, 1(3), 36–45.
- Klein, A. (2015). No Child Left Behind: An Overview. *Education Week*. Retrieved from <https://www.edweek.org/ew/section/multimedia/no-child-left-behind-overview-definition-summary.html>.
- Klenowski, V. (2011). Assessment for learning in the accountability era: Queensland, Australia. *Studies in Educational Evaluation*, 37(1), 78–83. <https://doi.org/10.1016/j.stueduc.2011.03.003>.
- Koh, K., & Luke, A. (2009). Authentic and conventional assessment in Singapore schools: an empirical study of teacher assignments and student work. *Assessment in Education: Principles, Policy & Practice*, 16(3), 291–318. <https://doi.org/10.1080/09695940903319703>.
- KSD. (2016, March 30). Menyikapi UN Pelajaran Tambahan Dimulai Pukul 06.00. *Berita Klaten*. Retrieved from <https://beritaklaten.com/30/03/2016/menyikapi-un-tambahan-pelajaran-dimulai-pukul-06-00/>.
- Lee, H. L. (2004). Our Future of Opportunity and Promise (National Rally Speech).
- Lee, J., & Reeves, T. (2012). Revisiting the Impact of NCLB High-Stakes School Accountability, Capacity, and Resources: State NAEP 1990 – 2009 Reading and Math Achievement Gaps and Trends. *Educational Evaluation and Policy Analysis*, 34(2), 209–231. <https://doi.org/10.3102/0162373711431604>.

- Lee, V. E., & Burkam, D. T. (2002). *Inequality at the Starting Gate: Social Background Differences in Achievement as Children Begin School*. Washington DC: Economic Policy Institute.
- Looney, A. (2014). Assessment and the Reform of Education Systems From Good News to Policy Technology. In C. Wyatt-Smith, V. Klenowski, & P. Colbert (Eds.), *Designing Assessment for Quality Learning: The Enabling Power of Assessment* (pp. 233–248). Dordrecht: Springer Science+Business Media.
- Masters, G. N. (2013). *Reforming Educational Assessment: Imperatives, principles and challenges*. Melbourne: Australian Council for Educational Research.
- Masters, G. N. (2014). Assessment: Getting to the essence, (1), 1–6.
- Moore, M. (2013, June 20). Riot after Chinese teachers try to stop pupils cheating. *The Telegraph*. Retrieved from <http://www.telegraph.co.uk/news/worldnews/asia/china/10132391/Riot-after-Chinese-teachers-try-to-stop-pupils-cheating.html>.
- Morris, P. (2002). Promoting curriculum reforms in the context of a political transition: an analysis of Hong Kong's experience. *Journal of Educational Policy*, 17(1), 13–28. <https://doi.org/10.1080/02680930110100036>.
- OECD. (2016). *PISA 2015 Assessment and Analytical Framework: Science, Reading, Mathematic and Financial Literacy*. Paris: OECD Publishing. <https://doi.org/10.1787/9789264255425-en>.
- OECD. (2018). *Equity in Education: Breaking Down Barriers to Social Mobility*. Paris: PISA, OECD Publishing.
- Osunsami, S., & Forer, B. (2011, July 7). Atlanta Cheating: 178 Teachers and Administrators Changed Answers to Increase Test Scores. ABC News. Retrieved from <https://abcnews.go.com/US/atlanta-cheating-178-teachers-administrators-changed-answers-increase/story?id=14013113>.
- Polesel, J., Rice, S., & Dulfer, N. (2013). The impact of high-stakes testing on curriculum and pedagogy: a teacher perspective from Australia. *Journal of Education Policy*, 29(5), 640–657. <https://doi.org/10.1080/02680939.2013.865082>.
- Puslitbang Sindo. (2018, March 29). Hadapi Ujian Nasional, Siswa Persiapkan 5 Langkah Ini. *Sindonews.Com*. Retrieved from <https://nasional.sindonews.com/read/1293437/144/hadapi-ujian-nasional-siswa-persiapkan-5-langkah-ini-1522236324#>.
- Roksa, J., & Potter, D. (2011). Parenting and academic achievement: Intergenerational transmission of educational advantage. *Sociology of Education*, 84(4), 299–321. <https://doi.org/10.1177/0038040711417013>.
- Schmidt, W. H., Burroughs, N. A., Zoido, P., & Houang, R. T. (2015). The Role of Schooling in Perpetuating Educational Inequality. *Educational Researcher*, 44(7), 371–386. <https://doi.org/10.3102/0013189X15603982>.
- Schoen, L., & Fusarelli, L. D. (2008). Innovation, NCLB, and the Fear Factor: The Challenge of Leading 21st-Century Schools in an Era of Accountability. *Educational Policy*, 22(1), 181–203.
- Sellen, R., Chong, K., & Tay, C. (2006). Assessment shifts in the Singapore education system. In *Annual IAEA Conference*. Singapore.
- Shihab, N. (2019, August 12). Kita yang Berpihak pada Anak, Agenda Prioritas Pendidikan Indonesia. *Kumparan*. Retrieved from <https://kumparan.com/najelaa-shihab/kita-yang-berpihak-pada-anak-agenda-prioritas-pendidikan-indonesia-1reTCNDaSV3>.
- Smith, J. M., & Kovacs, P. E. (2011). The impact of standards-based reform on teachers: the case of 'No Child Left Behind.' *Teachers and Teaching: Theory and Practice*, 17(2), 201–225. <https://doi.org/10.1080/13540602.2011.539802>.

- Subkhan, E. (2019). Mengapa adanya jasa bimbel bisa sulitkan pemerintah ketahui kualitas pembelajaran yang sebenarnya di sekolah. *The Conversation*. Retrieved from <https://theconversation.com/mengapa-adanya-jasa-bimbel-bisa-sulitkan-pemerintah-ketahui-kualitas-pembelajaran-yang-sebenarnya-di-sekolah-115012>.
- Sujatmiko. (2014, April 29). Sindikat Kunci Jawaban UN Ditangkap di Lamongan. *Tempo*. Co. Retrieved from <https://nasional.tempo.co/read/574108/sindikat-kunci-jawaban-un-ditangkap-di-lamongan/full&view=ok>.
- Supovitz, J. (2009). Can high stakes testing leverage educational improvement? Prospects from the last decade of testing and accountability reform. *Journal of Educational Change*, 10, 211–227. <https://doi.org/10.1007/s10833-009-9105-2>.
- Swain, K., Pendergast, D., & Cumming, J. (2018). Student experiences of NAPLAN: sharing insights from two school sites. *The Australian Educational Researcher*. <https://doi.org/10.1007/s13384-017-0256-5>.
- Tan, K. (2011). Assessment for learning reform in Singapore – quality, sustainable or threshold? In R. Berry & B. Adamson (Eds.), *Assessment Reform in Education: Policy and Practice* (pp. 75–87). Dordrecht: Springer Science+Business Media.
- Thompson, G. (2013). NAPLAN, MySchool and Accountability: Teacher perceptions of the effects of testing. *The International Education Journal: Comparative Perspectives*, 12(2), 62–84.
- Towles-reeves, E., Garrett, B., & Burdette, P. J. (2006). Validation of large-scale alternate assessment systems and their influence on instruction - What are the consequences? *Assessment for Effective Intervention*, 31(3), 45–57.
- UNESCO Institute for Statistics (2018). The impact of large scale learning assessment.
- Wurinanda, I. (2016, April 10). Beragam Kecurangan UN Terulang Setiap Tahun. *Okezone News*. Retrieved from <https://news.okezone.com/read/2016/04/09/65/1358480/beragam-kecurangan-un-terulang-setiap-tahun>.
- Wyn, J., Turnbull, M., & Grimshaw, L. (2014). *The Experience of Education: The impacts of high stakes testing on school students and their families. A qualitative study*. Sydney.
- Yu, L., & Suen, H. K. (2005). Historical and contemporary exam-driven education fever in China. *KEDI Journal of Educational Policy*, 2(1), 17–33.
- Zhao, Y. (2015). *Lessons that matter: What should we learn from Asia's school systems? Mitchell Policy Paper* (Vol. 4). Mitchell Institute.
- Zwick, R., & Green, J. G. (2007). New Perspectives on the Correlation of SAT Scores, High School Grades, and Socioeconomic Factors. *Journal of Educational Measurement*, 44(1), 23–45.

