

PENGUKURAN DALAM PENELITIAN PENDIDIKAN SENI

Hanna Sri Mudjilah

Pendahuluan

Pendidikan di Indonesia saat ini telah memberlakukan system pendidikan yang mengacu pada Kurikulum Tingkat Satuan Pendidikan (KTSP). Kurikulum tersebut memberikan kesempatan pada setiap satuan pendidikan untuk dapat mengembangkan materi, metode, sarana prasarana, bahkan sampai pada evaluasi, sesuai dengan kemampuan yang dapat dipenuhi oleh masing-masing satuan pendidikan. Kurikulum tersebut hanya memberikan batasan-batasan atau standar minimal yang harus dipenuhi, yaitu standar kompetensi dan kompetensi dasar, sehingga pembelajaran dapat berlangsung sebagaimana seharusnya.

Hal ini dapat menimbulkan berbagai macam kemungkinan variasi, baik materi, metode, maupun evaluasi yang dilakukan pada tiap satuan pendidikan. Adanya penetapan Undang-undang Nomor 20 tahun 2003, tentang Sistem Pendidikan Nasional dan Peraturan Pemerintah Nomor 19 tahun 2005, tentang Standar Nasional Pendidikan akan membawa dampak pada system penyelenggaraan pendidikan termasuk pengembangan dan pelaksanaan kurikulum. Kebijakan pemerintah tersebut berimplikasi kepada setiap satuan pendidikan agar dapat mengembangkan Kurikulum Tingkat Satuan Pendidikan, dengan mengacu pada Standar Nasional Pendidikan.

Permasalahan yang kemudian muncul adalah bahwa dalam melaksanakan KTSP tersebut, termasuk system penilaian dan evaluasinya, banyak dijumpai pendidik yang masih mengalami kesulitan untuk menyusun tes dan mengembangkan butir soal yang valid dan reliable. Hal ini sangat dirasakan khususnya bagi pendidik seni, yang sampai saat ini pun masih belum terdapat adanya rambu-rambu penilaian seni.

Berdasarkan beberapa informasi dan observasi yang diperoleh dari para guru, orangtua murid, maupun siswa, masih banyak pendidik yang belum melaksanakan penilaian (*assessment*) dengan baik terhadap anak didiknya, terutama pada mata pelajaran seni musik. Hal ini dapat menurunkan tingkat motivasi anak didik dalam belajar. Oleh sebab itu, maka ada himbuan bagi para pendidik untuk dapat memberikan penilaian terhadap prestasi atau kemampuan

anak didik sesuai dengan kaidah-kaidah penilaian yang sesungguhnya, dan dapat dipertanggungjawabkan.

Suatu lembaga pendidikan baik formal maupun informal dapat dikatakan telah melakukan pekerjaannya dengan baik apabila dapat membuktikan bahwa anak didiknya memperoleh kemampuan yang baru sebelum mengikuti suatu pembelajaran. Hal ini ditegaskan oleh layanan yang bergerak di bidang pengujian pendidikan, yaitu *Educational Testing Service*, yang menyatakan bahwa tanpa pengukuran yang reliable dari apa yang telah dipelajari oleh siswa, hal ini sangat tidak mungkin untuk menyatakan bahwa sekolah tersebut telah melakukan suatu pekerjaan yang baik. Kazin dan Payne (2009), menyatakan bahwa:

1. *Without assessing student learning outcomes, there is no reliable way to measure and demonstrate an institution's*
2. *Educational quality. Because accrediting bodies increasingly are seeking evidence of student-learning outcomes, more governing boards and top administrators will need to take assessment seriously in the future.*
3. *The process of creating a "culture of evidence" encourages colleges and universities to reflect on what aspirations they have for their students and then to generate evidence concerning how well they are meeting those goals.*

Mutu pendidikan di Indonesia menurut *The Third International Mathematics and Science Study (TIMSS)* tahun 1999, berada pada urutan ke-32 untuk IPA, dan ke-34 untuk Matematika dari 38 negara peserta. Bahkan jika dibandingkan dengan hasil survey dari system *Polotical and Economic Risk Consultancy (PERC)*, menyimpulkan bahwa system pendidikan Indonesia berada pada urutan terakhir dari 12 negara.

Terkait dengan kondisi pendidikan di Indonesia yang berada pada posisi yang kurang menggembirakan, maka dilakukan pembaruan dan penyempurnaan secara menyeluruh, sehingga dilakukan penyempurnaan terhadap kurikulum nasional untuk pendidikan dasar sampai menengah. Perubahan kurikulum tersebut membuat adanya pergeseran paradigma pendidikan di Indonesia, yaitu dari pendekatan pendidikan yang berorientasi pada masukan, menjadi pendekatan pendidikan yang berorientasi pada hasil atau standard. (Hayat: 2004). Pergeseran paradigma tersebut menuntut adanya suatu standard kompetensi yang harus dicapai oleh peserta didik. Lebih jauh dikatakan bahwa, dalam konteks pendidikan, standard diperlukan sebagai acuan minimal yang

harus dipenuhi oleh lulusan dari suatu lembaga pendidikan sehingga setiap calon lulusan dinilai apakah yang bersangkutan telah memenuhi standard minimal yang telah ditetapkan.

Untuk memperoleh informasi tentang prestasi belajar peserta didik, diperlukan tahapan yang disebut dengan pengukuran (*measurement*), penilaian (*assessment*), ataupun evaluasi. Evaluasi dilakukan setelah terlebih dahulu mengadakan pengukuran dan penilaian terhadap prestasi belajar peserta didik. Oleh karena evaluasi dilakukan untuk memperoleh hasil yang akurat terhadap prestasi belajar peserta didik, maka perlu dilakukan dengan benar.

Kenyataan saat ini, penilaian maupun evaluasi terhadap prestasi belajar peserta didik, khususnya pada mata pelajaran seni musik, masih belum memiliki standar yang benar. Penilaian yang dilakukan masih sering dipengaruhi oleh subjektivitas pendidik, sehingga seringkali diperoleh penilaian yang sulit untuk dipertanggungjawabkan.

Kurikulum Tingkat Satuan Pendidikan (KTSP) jenjang Pendidikan Dasar dan Menengah, seperti ditentukan dalam Peraturan Pemerintah Nomor 19 Tahun 2005, tentang Standar Nasional Pendidikan, memberi kesempatan setiap sekolah/madrasah untuk mengembangkan kurikulum tersebut berdasarkan pada Standar Kompetensi Lulusan (SKL) dan Standar Isi (SI), serta berpedoman pada panduan yang telah ditetapkan oleh Badan Standar Nasional Pendidikan (BSNP). Oleh karena setiap sekolah/madrasah diberikan kebebasan untuk mengembangkan kurikulumnya, maka sangat dimungkinkan terdapat berbagai macam kemungkinan model/variasi kurikulum yang ada, walaupun dikembangkan berdasarkan pada standar-standar seperti yang telah ditentukan.

Namun demikian, adanya Peraturan Pemerintah No. 19 Tahun 2005 tentang Standar Nasional Pendidikan yang menegaskan bahwa Standar Nasional Pendidikan adalah kriteria minimal tentang sistem pendidikan di seluruh wilayah hukum Negara Kesatuan Republik Indonesia, diharapkan dapat meminimalisir terjadinya perbedaan pada pembelajaran di sekolah-sekolah, karena ketidakmampuan dari para pengambil kebijakan di sekolah terhadap standar pendidikannya. BSNP sebagai sebuah lembaga independen, bertugas membantu Menteri dalam mengembangkan, memantau, dan mengendalikan standar nasional pendidikan, terhadap: (1) Standar Isi, (2) Standar Kompetensi Lulusan, (3) Standar Proses, (4) Standar Tenaga Pendidikan dan Tenaga

Kependidikan, (5) Standar Sarana dan Prasarana, (6) Standar Pengelolaan, (7) Standar Pembiayaan, dan (8) Standar Penilaian Pendidikan.

Penelitian Pendidikan Seni

Penelitian pada awalnya hanya dilakukan pada bidang eksakta dan dilakukan di dalam laboratorium. Akan tetapi, penelitian-penelitian sosial berkembang belakangan ini dengan sangat pesatnya, termasuk di dalamnya penelitian seni. Indonesia, sampai saat ini masih belum banyak melaksanakan penelitian di bidang seni, seperti layaknya di Negara-negara maju. Penelitian Seni merupakan penelitian yang dilakukan terhadap berbagai permasalahan tentang seni, baik seni rupa, seni tari, seni musik, maupun cabang seni yang lain.

Penelitian pendidikan seni sebagai bagian dalam penelitian seni, ternyata memiliki berbagai permasalahan yang masih dapat diteliti. Hal ini dikarenakan penelitian pendidikan seni masih sangat langka, bila dibandingkan dengan penelitian-penelitian di bidang lain. Masih sedikitnya ilmuwan, para guru, para pendidik, yang belum meluangkan waktunya dengan baik untuk melakukan sebuah penelitian yang ditujukan bagi perbaikan ilmu maupun perbaikan pembelajaran.

Banyak hal yang dapat dikaji dari permasalahan-permasalahan dalam pembelajaran seni, khususnya seni musik. Paling tidak, jika ditinjau dari ketiga ranah yang dicetuskan oleh seorang ahli pendidikan, Bloom, yaitu ranah kognitif, afektif, dan psikomotor, sangat banyak permasalahan yang dapat diangkat menjadi sebuah penelitian. Tiap ranah memiliki ciri khas masing-masing. Dalam sebuah penelitian pembelajaran seni, diperlukan adanya pengambilan data terhadap subjek penelitian, yang dalam hal ini adalah peserta didik.

Untuk mengungkap kemampuan peserta didik pada ranah kognitif sudah sering dilakukan oleh para pendidik, namun untuk kedua ranah yang lain, yaitu afektif dan psikomotor masih jarang dilakukan, karena membutuhkan seperangkat instrumen atau alat pengumpul yang perlu dipersiapkan terlebih dahulu. Instrumen dalam sebuah penelitian sangat menentukan kualitas penelitian itu sendiri, sehingga apabila instrumen yang dihasilkan itu baik, maka kualitas hasil dari penelitian tersebut akan baik dan dapat dipercaya. Sedangkan jika kualitas instrumen itu tidak baik, maka hasil penelitian tersebut menjadi tidak

baik dan tidak dapat dipercaya. Hal ini akan dapat berakibat pada pengambilan keputusan kebijakan dari hasil penelitian.

Instrumen dalam sebuah penelitian pendidikan (sosial) dapat berbentuk tes, dan angket, yang dikerjakan oleh subjek penelitian, maupun lembar observasi yang dikerjakan oleh *rater* (peneliti). Sedangkan kualitas suatu instrumen penelitian dapat diperoleh dengan cara mengukur reliabilitas dan validitas dari instrumen tersebut.

Untuk mendapatkan sebuah instrumen yang baik dan dapat dipertanggungjawabkan, perlulah memenuhi syarat-syarat penting sehingga instrumen tersebut dapat berfungsi sebagaimana seharusnya. Ada tiga syarat penting yang harus dipenuhi, yaitu (1) reliabilitas, (2) validitas, dan (3) visibilitas.

Instrumen tes yang baik akan memiliki tingkat reliabilitas yang tinggi. Instrumen tes yang baik akan dapat memberi hasil yang sama walaupun dilakukan oleh tester yang berbeda, ataupun diskor oleh orang yang berbeda, bentuk instrumen tes yang diberikan berbeda, dan orang yang sama melakukan tes pada waktu yang berbeda, hasilnya harus tetap sama. Terkait dengan reliabilitas, Hadi (2000:36), menyatakan bahwa terdapat tiga macam pendekatan reliabilitas, yaitu: (1) pendekatan tes ulang (*tes retest*); (2) pendekatan bentuk paralel (*parallel-forms*); (3) dan pendekatan konsistensi internal (*internal consistency*).

Syarat yang lain dari sebuah instrumen yang baik adalah validitas. Artinya bahwa instrumen tersebut harus dapat mengukur apa yang seharusnya diukur. Cureton menyatakan bahwa "*The essential question of test validity is how well a test does the job it is employed to do*" (Djaali, 2008: 49). Menurut Hadi (2000:45), ada tiga kategori validitas tes, yaitu (1) validitas isi (*content validity*); dan (2) validitas konstrak (*construct validity*), dan validitas berdasar kriteria (*criterion-related validity*). Validitas isi dapat diestimasi melalui pengujian terhadap isi tes atau dapat melalui *professional judgment*. Validitas isi terbagi menjadi dua tipe, yaitu *face validity*, dan *logical validity (sampling validity)*. Validitas konstrak adalah validitas yang mempermasalahkan seberapa jauh item-item tes mampu mengukur apa yang benar-benar hendak diukur sesuai dengan konsep khusus atau definisi konseptual yang telah ditetapkan (Djaali: 2008, 51).

Visibilitas merupakan salah satu syarat lain dari sebuah instrumen penelitian yang baik. Instrumen yang baik perlu ditinjau dari sisi penampilan,

apakah instrumen tersebut secara visual mudah untuk dipahami oleh subjek penelitian. Apabila instrumen penelitian tersebut sulit dipahami oleh subjek penelitian, maka akan terjadi kesalahan dalam memberikan respons dari apa yang dimaksud dalam instrumen tersebut. Penampilan sebuah instrumen penelitian haruslah menarik, simple, dan mudah dimengerti.

Penelitian sangat diperlukan bagi para pendidik, ilmuwan, maupun para pengambil kebijakan. Penelitian di bidang pembelajaran seni musik juga sangat diperlukan karena akan dapat diperoleh informasi apakah selama ini penilaian yang dilakukan oleh pendidik sudah benar, sesuai dengan penilaian yang sebenarnya. Atau, sejauhmanakah dampak dari hasil penilaian pendidik berdampak pada prestasi peserta didik.

Apabila terjadi kesalahan atau ketidaktepatan penilaian karena kurang tepatnya pengukuran yang dilakukan oleh pendidik, akan dapat berpengaruh secara signifikan pada semangat dan motivasi peserta didik dalam mengikuti proses pembelajarannya. Seorang peserta didik yang seharusnya mendapatkan nilai tinggi, tetapi karena kesalahan pengukuran, maka kemungkinan dia akan mendapatkan nilai yang rendah. Hal ini sangat dikhawatirkan dapat berpengaruh pada prestasi, bahkan motivasi dari peserta didik.

Pertanyaan yang sering muncul adalah, apakah pengukuran dalam pembelajaran seni musik itu ada? Apakah musik dapat diukur? Beberapa pertanyaan yang lain masih sering bermunculan, karena ketidaktahuan tentang pemahaman pengukuran, penilaian, dan evaluasi.

Penelitian seni musik, khususnya di bidang pendidikan seni musik masih belum banyak dilakukan, terutama dalam penelitian kuantitatif. Hal ini sering dijadikan sebagai *momok*, karena kebanyakan para pelaku pendidik seni lebih sering menghindarkan dari hal-hal yang berbau numerik (angka-angka). Bahkan ada *issue* bahwa seni tidak dapat dinilai, dan apabila dipaksakan untuk dinilai, maka akan selalu diperoleh penilaian yang subjektif. Untuk menghindarkan dari semua ini, maka suatu penelitian haruslah dilakukan dengan profesional.

Pengukuran, Penilaian, dan Evaluasi Penelitian Seni

Para pendidik sering melupakan bahwa tugas utama selain mengajar, adalah memberikan evaluasi kepada anak didiknya. Evaluasi yang dilakukan oleh pendidik haruslah mencakup ketiga ranah, yaitu kognitif, afektif, dan

psikomotor. Penilaian terhadap ranah afektif dan psikomotor masih sering mengalami banyak hambatan. Demikian juga untuk memberikan evaluasi khususnya pada ranah afektif dan psikomotor masih sering dipengaruhi oleh subjektivitas yang tinggi, karena belum dilakukan seperti sebagaimana seharusnya.

Para pendidik sangat perlu untuk menguasai bagaimana cara mengukur dan menilai, khususnya untuk ranah afektif dan psikomotor, sehingga seluruh hasil evaluasi yang diberikan kepada anak didik dapat objektif dan dapat dipertanggungjawabkan. Apabila seorang pendidik seni (musik) ingin melakukan penelitian, maka pemahaman tentang pengukuran dan penilaian perlu dikuasai dengan baik, sehingga tidak terjadi kesalahan pemahaman terhadap kedua istilah tersebut.

Evaluasi dalam penelitian dimaknai sebagai suatu proses yang harus dilakukan untuk membuat suatu keputusan terhadap suatu hasil penelitian. Sebelum melakukan suatu evaluasi, perlu dilakukan terlebih dahulu pengukuran terhadap beberapa indikator penelitian yang dibutuhkan untuk mengungkap suatu permasalahan, kemudian barulah dilakukan suatu penilaian terhadap hasil pengukuran tersebut. Untuk menghindarkan subjektivitas dalam melakukan pengukuran, seorang pendidik perlu mempersiapkan instrumen pengukuran yang mampu mengukur apa yang hendak diukur. Setelah diperoleh hasil pengukuran tersebut barulah kemudian dilakukan penilaian, yang kemudian dapat dilakukan untuk memberikan evaluasi terhadap kemampuan anak didik.

Jenis Tes dalam Pendidikan

Istilah tes dan pengukuran, seringkali dipertukarkan satu sama lain. Pengertian tes menurut Brown (dalam Azwar 2000: 3) adalah prosedur yang sistematis guna mengukur sampel perilaku seseorang. Sedangkan Cronbach memberikan definisi tes sebagai "... a systemic procedure for observing a person/s behavior and describing it with the aid of a numerical scale or a category system" (Azwar 2000:3). Tyler dalam Azwar (2000:4) mengatakan bahwa pengukuran adalah "...assignment of numerals according to rules".

Pengertian Tes banyak dimaknai oleh beberapa tokoh seperti Cronbach (1984), bahwa tes adalah suatu prosedur yang sistematis untuk mengamati atau mendeskripsikan satu atau lebih karakteristik seseorang dengan menggunakan

standar numerik atau sistem kategori. Bruce (1978) memaknai tes dapat digunakan untuk mengukur banyaknya pengetahuan yang diperoleh individu dari suatu bahan pelajaran yang terbatas pada tingkat tertentu. Sedangkan menurut Norman (1976), memaknai tes sebagai salah satu prosedur evaluasi yang komprehensif, sistematis, dan objektif yang hasilnya dapat dijadikan sebagai dasar dalam pengambilan keputusan dalam proses pengajaran yang dilakukan oleh guru.

Ada beberapa jenis tes yang dikelompokkan berdasarkan pada tujuan dilakukannya tes. Adapun jenis-jenis tes tersebut adalah:

- Tes diagnostik, untuk mendapatkan informasi tentang konsep-konsep yang belum dipahami, maupun yang sudah dipahami.
- Tes formatif, untuk mengetahui keberhasilan proses pembelajaran.
- Tes sumatif, untuk menentukan keberhasilan belajar peserta didik untuk pelajaran tertentu.

Adapun fungsi diadakannya tes adalah (1) sebagai alat untuk mengukur prestasi belajar siswa, (2) sebagai motivator dalam suatu pembelajaran (Thorndike, 1991; Ebel, 1979), (3) sebagai upaya perbaikan kualitas pembelajaran. Sebagai upaya perbaikan kualitas pembelajaran dapat berbentuk tes penempatan, tes diagnostik, tes formatif, maupun tes sumatif.

Tes formatif dilakukan untuk mendapatkan umpan balik bagi usaha perbaikan kualitas pembelajaran dalam konteks kelas. Airasian (2005: 151), berpendapat bahwa *formative assessments are used to alter or improve instruction while it is still going on.*

Tes sumatif sebagai tes hasil belajar dilakukan setelah sekumpulan materi pelajaran atau satuan program pengajaran selesai diberikan. Tes sumatif dapat untuk menentukan kedudukan (rangking) tiap siswa, dapat memutuskan apakah seorang siswa dapat melanjutkan program pembelajarannya atau tidak, selain itu, tes sumatif dapat menginformasikan kemajuan siswa kepada pihak lain. Airasian (2005: 151), juga berpendapat bahwa *summative assessments are used to evaluate the outcomes of instruction and take the form of tests, projects, term papers, and final exams.* Grades, sebagai bentuk yang sering ditemukan pada *summative assessment.* (Middle States Commission on Higher Education. 2007: 27) www.msche.org.

Ada jenis tes lain yaitu *Official assessments*, yang disebutkan bahwa *official assessments are needed by the school bureaucracy for purposes such as pupil testing, grading, and placement* (Airasian, 2005: 151). Jenis tes ini dibutuhkan oleh birokrasi sekolah dengan tujuan seperti melakukan tes siswa, menggolongkan berdasarkan tingkat, dan mengelompokkan siswa. Dalam penilaian kelas, *official assessment* seringkali ditujukan pada *pupils' cognitive performance*, biasanya seberapa baik apa yang harus dipahami, telah dipelajari oleh siswa.

Tinjauan terhadap tes berdasarkan pada aspek psikis, dibedakan menjadi:

1. *Intelligency test* : tes inteligensi, untuk memprediksi tingkat kecerdasan seseorang
2. *Aptitude test* : tes kemampuan untuk mengungkap kemampuan dasar seseorang atau bakat khusus seseorang
3. *Attitude test* : tes sikap, untuk mengungkap pre-disposisi, atau kecenderungan merespon objek tes kepribadian, untuk mengungkap ciri
4. *Personality test* : khas seseorang tes hasil belajar, untuk mengungkap tingkat
5. *Achievement test* : pencapaian terhadap tujuan pembelajaran atau prestasi belajar

Tes psychology yang telah dikenal ada beberapa jenis, antara lain: tes IQ, EQ, SQ, dsb. Masing-masing tes tersebut telah dikembangkan dan memiliki standar yang sama, sehingga seorang yang telah mengikuti tes IQ di salah satu tempat, akan diakui juga hasil dari tes tersebut di tempat lain.

Ada beberapa pertanyaan jika dan bagaimana 'test' dapat digunakan untuk menyeleksi secara adil dan akurat? Sloboda (1985, 234-236), menyatakan bahwa:

1. *A test should only be used when there are no more direct signs of achievement to examine*
2. *Test results should be taken in conjunction with other evidence, formal and informal, where possible.*

3. *Testing should only be carried out with respect to a particular educational question to be answered at a particular time, not to provide a once-for-all statement about capacity or potential for achievement.*
4. *The test should be a valid measure of musical ability. Validity may be assessed in two ways. A test has 'face' or 'content' validity if it manifestly demands use of the ability it is purporting to test.*
5. *The test should be reasonably reliable. A test is reliable if the same subject achieves the same score when tested more than once.*

Instrumen Pengukuran

Instrumen pengukuran dalam penelitian di bidang seni musik, khususnya di Indonesia masih sangat kurang. Hal ini disebabkan karena banyaknya indikator yang diperlukan untuk mengukur baik ranah afektif maupun psikomotor sangat bervariasi dari satu instrumen dengan instrumen yang lain. Setelah instrumen pengukuran selesai disusun, masih diperlukan satu langkah untuk mengukur validitas dan reliabilitas instrumen tersebut.

Beberapa ahli berpendapat bahwa untuk menyusun sebuah instrumen pengukuran, diperlukan beberapa langkah berupa tahapan-tahapan yang harus dilalui, untuk akhirnya didapat sebuah instrumen yang dapat digunakan untuk pengambilan data penelitian. Mardapi (2008: 108), menyatakan bahwa ada 10 langkah yang harus diikuti dalam mengembangkan instrumen afektif, yaitu:

1. menentukan spesifikasi instrumen
2. menulis instrumen
3. menentukan skala instrumen
4. menentukan system penskoran
5. menelaah instrumen
6. melakukan ujicoba
7. menganalisis instrumen
8. merakit instrumen
9. melaksanakan penskoran
10. menafsirkan hasil pengukuran

Skala pengukuran dari tes tersebut dapat dilakukan dengan berbagai jenis skala pengukuran, antara lain dengan skala Likert maupun Thurstone.

Ada kalanya tes-tes musik yang telah ada tidak mampu untuk diberikan pada anak-anak kecil dan anak-anak dengan kebutuhan khusus, karena mereka mempunyai kebutuhan yang berbeda terhadap instruksi verbal dan respon verbal dari orang dewasa. Sebagai contoh, untuk memenuhi kebutuhan tersebut,

Matsuyama mengembangkan sebuah tes untuk menilai respon musical dari anak-anak kecil. Tes tersebut dinamakan *Non-Verbal Measurement of the Musical Responsiveness of Children* (Non-Verbal MMRC). Tes ini terdiri dari dua bagian, yaitu:

1. *Rhythm section:*

- Single quarter note;
- Sequential two quarter notes;
- Sequential three quarter notes;
- Single quarter note—Sequential two eighth notes—Single quarter note;
- Triplet—single quarter note;
- Dotted notes—single quarter note;
- Sequential four sixteenth notes—single quarter note.

2. *Melody section:*

- Single quarter note
- Major 2nd (turun);
- Major 3rd (turun);
- Contour scheme of major 2nd (turun-naik);
- Contour scheme of major 3rd (turun-naik);
- Contour scheme of major 3rd (circulation).

(<http://www.medscimonit.com/dulltxt.php?IDMAN=7268>).

Pemahaman tentang *assessment* dan *measurement* dijelaskan oleh Guion bahwa *assessment* adalah suatu istilah yang lebih luas maknanya dibanding dengan *measurement*. *Measurement* merupakan kasus yang special dari *assessment*. Lebih lanjut dikatakan bahwa:

....It based on a more defined scale along which scores can be ordered with relatively fine gradations (e.g., measurement of mechanical ability). Measurement seeks precision. Incontrast, many other assessment procedures are ad hoc or used for specific practical purposes where precision is not useful or perhaps not possible. (Guion, 2006: 115)

Analisis Instrumen diperlukan untuk melihat apakah instrumen tersebut mampu menghasilkan variasi jawaban. Jika nilai variasi jawaban semakin besar, maka instrumen tersebut akan semakin baik. Sebaliknya, jika variasi jawaban itu sedikit (kecil), maka dikatakan bahwa instrumen tersebut kurang baik atau tidak mampu membedakan antara kemampuan atau *ability* dari *examinee*. Hal ini dikarenakan instrumen yang digunakan untuk mengumpulkan data harus mampu memberikan informasi yang akurat.

Setelah terbukti bahwa instrumen tersebut memiliki nilai variasi yang besar, maka untuk mengukur keterandalan instrumen dilakukan dengan formula dari Cronbach-alpha. Jika nilai indeks alpha sama dengan atau lebih besar dari 0.70, maka instrumen tersebut tergolong baik. (Mardapi, 2008: 122).

Model Pengukuran

Teori tes modern, yang dikenal dengan item response theory (IRT), saat ini mendukung classical test theory (CTT), dimana IRT merupakan model pengukuran untuk data tes, *scoring test*, dan *reporting test*. Kemungkinan untuk melakukan pengujian dengan *computerized adaptive test* membuat Lord mengembangkan IRT pada tahun 1960 dan 1970an. Dikatakan oleh Lord (1980), bahwa model-model pengukuran diperlukan untuk membangun seleksi optimal dari item test dan skor estimasi dari *examinee*. Model-model statistik diperlukan sehingga skor *examinee* dapat dibandingkan satu dengan yang lain, walaupun *examinee* dilakukan dengan subtes yang berbeda dari item tes. Teori tes modern dikembangkan untuk memenuhi kebutuhan tersebut.

(Hambleton, <http://apm/sagepub.com/cgi/content/refs/24/4/291>).

Samejima (1969) melanjutkan pekerjaan dari Lord dengan memfokuskan pada model-model IRT pada kelas untuk data skor politomus (seperti, *Graded Response Model*). Brennan mengatakan bahwa penggunaan teori *Generalizability* adalah untuk menguji faktor-faktor, seperti *tasks, raters, methods, occasions*, dan interaksinya dan peran pada kinerja *examinee*. Robert Brennan merupakan salah satu tokoh dalam mengaplikasi *Generalizability* pada *Performance Assessments (PA)*. (Hambleton, <http://apm/sagepub.com/cgi/content/refs/24/4/291>).

Penelitian Pendidikan Seni, khususnya pada seni pertunjukan, seringkali menghadapi permasalahan dalam pengambilan data jika ingin diperoleh data penelitian berupa performance. Untuk mengatasi permasalahan tersebut, seorang peneliti haruslah mempersiapkan terlebih dahulu indikator dari variabel penelitian yang akan digunakan sebagai data penelitian. Langkah awal yang harus ditempuh adalah menyusun indikator yang ingin diungkap dalam sebuah penelitian, kemudian dari masing-masing indikator tersebut diberikan penjelasan dan batasan-batasan dalam bentuk rubrik, sehingga pengambil data, dalam hal

ini para rater tidak akan memaknai sebuah indikator berbeda satu dengan yang lain.

Penutup

Saat ini, penelitian di bidang pendidikan khususnya pendidikan musik sangat terbatas. Hal ini disebabkan karena beberapa diantara para peneliti, masih banyak mengalami hambatan. Hambatan-hambatan yang selama ini dijumpai adalah karena seni (musik), telah di *judge* bahwa seni sulit untuk diukur dan dinilai. Memang selama ini masih terjadi pemaknaan yang berbeda antara ahli yang satu dengan yang lain.

Artikel ini ingin menjembatani adanya pengertian maupun pemahaman yang berbeda. Pengukuran dan penilaian dalam suatu evaluasi pendidikan musik sangat mungkin dilakukan jika para pengambil keputusan memaknai pengukuran dan penilaian dengan benar.

Harapan dari artikel ini agar para pelaku pendidikan akan mendapatkan manfaat positif dalam penelitian pendidikan seni (musik), yang kemudian dapat diterapkan pada institusi terkait.

DAFTAR PUSTAKA

- Alexander, Bryant K, et.al. 2005. *Performance Theories in Education*. Mahwah: Lawrence Erlbaum Associates, Publishers.
- American Educational Research Association. 1999. *Standards for Educational and Psychological Testing*. Washington
- Azwar, Saifuddin. 2000. *Reliabilitas dan Validitas*. Yogyakarta: Pustaka Pelajar.
- Blalock, Hubert M. 1989. *Conceptualization and Measurement in The Social Sciences*. Newbury Park: Sage Publications.
- Brennan, Robert L. 2000. *Performance Assessments from the Perspective of Generalizability Theory*. Applied Psychological Measurement vol. 24: 339-353. <http://apm.sageub.com/cgi/content/refs/24/4/339>
- Caruso, David R. Mayer, John D. and Salovey, Peter. 2002. *Relation of an Ability Measure of Emotional Intelligence to Personality*. Journal of Personality Assessment vol. 79: 306-320. Lawrence Erlbaum Associates, Inc.
- Choksy, Lois. 1981. *The Kodaly Context: Creating an Environment for Musical Learning*. Englewood Cliffs: Prentice-Hall, Inc.
- Cross, Tracy.L, et.all. 2008. *The Psychology of Gifted Adolescents as Measured by the MMP-A*. Gifted Child Quarterly vol. 52: 326-339. <http://gcq.sagepub.com/cgi/content/refs/52/4/326>
- Djaali,H dan Muljono, Pudji. 2008. *Pengukuran Dalam Bidang Pendidikan*. Jakarta: PT Grasindo.
- Djohan. 2005. *Psikologi Musik*. Yogyakarta: Penerbit Buku Baik.
- _____. 2006. *Terapi Musik*. Yogyakarta: Penerbit Galang Press.
- Earl, Lorna. 2003. *Assessment as Learning: Using Classroom Assessment to Maximise Student Learning*. Thousand Oaks, CA, Corwin Press.
- Edwards, Alistair, DN. et.al. *Development of a standard test of musical ability for participants in auditory interface testing*. <http://www.icad.org/websitev2.0/conferences/ICAD2000/PDFS/Edwards.pdf>.
- Frith, DS, Macintosh, HG. 1988. *A Teacher's Guide to Assessment*. Glasgow: Bell and Baik Ltd.
- Gregory, Robert J. 2000. *Psychological Testing*.Third Edition. Needham Heights: Allyn & Bacon, Inc.

- Gronlund, Norman E. 1982. *Constructing Achievement Tests*. Third Edition. Englewood Cliffs: Prentice-Hall, Inc.
- Guion, Robert M. 2006. *Essentials of Personal Assessment and Selection*. Mahwah: Lawrence Erlbaum Associates, Publishers.
- Hallam, Susan. 2006. *Conceptions of musical ability*. <http://www.marcocosta.it/icmpc2006/pdfs/126.pdf>
- Haroutounian, Joanne. *Identification of the Musically Talented Student: The Assessment of Musical Potential and Musical Performance*.
- Harrison, Carole S. 1987. *The Validity of Musical Aptitude Profile for Predicting Grades in Freshman Music Theory*. Educational and Psychological Measurement vol. 47: 477-482). <http://epm.sagepub.com>
- Holsomback, J. Richard, Jr. 2001. *Evaluating the Relationship Between Musical Aptitude and Standardized Achievement Test Scores of Beginner Instrumental Music Students*. Texas Music Education Research. (1-8).
- Lawther, John D. 1977. *The Learning And Performance of Physical Skills*. Second Edition. Englewood Cliffs: Prentice-Hall, Inc.
- Lee, Donghyuck and Pfeiffer, Steven I. 2006. *The Reliability and Validity of a Korean-Translated Version of the Gifted Rating Scales*. Journal of Psychological Assessment vol. 24: 210-224. <http://jpa.sagepub.com/cgi/content/refs/24/3/210>
- Li, Huijun, et.all. 2008. *Validation of the Gifted Rating Scales School Form in China*. Gifted Child Quarterly vol. 52: 160-169. <http://gcq.sagepub.com/cgi/content/refs/52/2/160>
- Mardapi, Djemari. 2008. *Teknik Penyusunan Instrumen Tes dan Nontes*. Jogjakarta: Mitra Cendikia Press.
- Maryland State Improvement Grant. 2003. *Performance Assessments: A Resource For Special Education Teacher Educators in Maryland*. Baltimore: Division of Certification and Accreditation
- Napoles, Jessica and Madsen, Clifford. K. 2008. *Measuring emotional response to usic within a classroom setting*. International Journal of Music Education vol.26: 63-71. <http://ijm.sagepub.com/cgi/content/refs/26/1/63>
- Osborne, Margaret S. et.all. 2005. *Assessment of music performance anxiety in late childhood: a validation study of the Music Performance Anxiety Inventory for Adolescents (MPAI-A)*. Sydney: Australian Centre for Applied Research in Music Performance (ACARMP).

- Petscher, Yaacov and Li, Huijun. 2008. *Measurement Invariance of the Chinese Gifted Rating Scales: Teacher and Parent Forms*. Journal of Psychological Assessment vol. 26: 274-286.
<http://jpa.sagepub.com/cgi/content/refs/26/3/274>
- Portowitz, Adena and Klein, Prina S. 2007. *MISC-MUSIC: a music program to enhance cognitive processing among children with learning difficulties*. International Journal of Music Education vol. 25: 259-271.
<http://ijm.sagepub.com/cgi/content/refs/25/3/259>
- Sloboda, John. A. 1990. *The Musical Mind – The Cognitive Psychology of Music*. New York: Oxford University Press.
- Sugiyono. 2006. *Metode Penelitian Kuantitatif Kualitatif dan R & D*. Bandung: Penerbit Alfabeta.
- Torre, Jimmy de la. 2008. *Multidimensional Scoring of Abilities: The Ordered Polytomous Response Case*. Applied Psychological Measurement vol. 32: 355-370. <http://apm.sagepub.com/cgi/content/refs/32/5/355>
- Williams, Thomas.O, et.all. 2002. *Confirmatory Factor Analysis of an Instrument Designed to Measure Affective and Cognitive Arousal*. Educational and Psychological Measurement vo. 62: 264-283.
<http://epm.sagepub.com/cgi/content/refs/62/2/64>