

PANDUAN PENJAMINAN MUTU ASESMEN DAN EVALUASI PEMBELAJARAN DARING

**DIREKTORAT PEMBELAJARAN DAN KEMAHASISWAAAN
DIREKTORAT JENDERAL PENDIDIKAN TINGGI
KEMENTERIAN PENDIDIKAN DAN KEBUDAYAAN
2014**

Sambutan

Terbitnya 4 (empat) buku Panduan Penjaminan Mutu yang terdiri atas: Panduan Penjaminan Mutu Konten, Panduan Penjaminan Mutu Proses Pembelajaran, Panduan Penjaminan Mutu Asesmen dan Evaluasi, serta Panduan Penjaminan Mutu Sistem ini mengawali dicanangkannya Pembelajaran Daring Indonesia Terbuka dan Terpadu (PDITT) yang merupakan upaya bersama perguruan-perguruan tinggi di Indonesia untuk bahu-membahu menyajikan layanan pendidikan tinggi berkualitas bagi seluruh masyarakat Indonesia.

Sebagai bagian dari sistem pendidikan jarak jauh di Indonesia, PDITT dimaksudkan sebagai cara untuk melayani masyarakat secara luas dan masif dengan tidak mengorbankan kualitas. Oleh karena itu, diterbitkannya buku panduan penjaminan mutu ini juga dimaksudkan sebagai upaya pengelola PDITT untuk secara transparan dan akuntabel mempertanggungjawabkan pengelolaan program penting ini.

Pencanangan PDITT yang diawali dengan memberikan kepercayaan kepada 5 perguruan tinggi yakni Institut Teknologi Bandung, Institut Teknologi Sepuluh Nopember, Universitas Indonesia, Universitas Gadjah Mada, Universitas Bina Nusantara, dan Asosiasi Perguruan Tinggi Komputer untuk menyajikan mata kuliah yang dapat diambil oleh seluruh mahasiswa maupun publik yang berminat. Setelah pencanangan dan rintisan penyelenggaraan dilakukan, tahapan berikutnya adalah memberikan kesempatan kepada dosen dan pakar di seluruh Indonesia berontribusi menyediakan mata kuliah maupun materi pembelajaran bermutu bagi seluruh mahasiswa dan publik Indonesia. Dalam konteks menjamin mutu seluruh proses pembelajaran inilah, keempat buku Panduan Penjaminan Mutu ini digunakan.

Kami mengucapkan terima kasih kepada Tim Pengembang PDITT yang telah dengan tekun membantu Direktorat Pembelajaran dan Kemahasiswaan melakukan berbagai persiapan, termasuk di dalamnya menyelesaikan keempat buku panduan ini. Harapan kami, semua upaya yang kita lakukan ini bermanfaat bagi bangsa Indonesia terutama masyarakat perguruan tinggi.

Jakarta, September 2014
Direktur Pembelajaran dan Kemahasiswaan

Illah Sailah
NIP 19580521 198211 2001

1. Pendahuluan	4
1.1 Latarbelakang	4
1.2 Tujuan	5
2. Pengertian dan Ruang Lingkup Asesmen dan Evaluasi Program Pembelajaran Daring	6
2.1 Pengertian asesmen dan evaluasi	6
2.2 Asesmen dalam pembelajaran daring	7
2.3 Prinsip Asesmen Pembelajaran daring	8
2.4 Evaluasi Program Pembelajaran daring	11
2.5 Ruang lingkup asesmen dan evaluasi pembelajaran daring	16
2.6 Mutu dan rancangan asesmen daring	16
3. <i>Standart Operating Procedure</i>	22
4. Daftar Pustaka	24

1. Pendahuluan

Latar Belakang

Perkembangan teknologi dan komunikasi pada saat ini sangat memungkinkan untuk dimanfaatkan memindahkan proses pembelajaran dari dalam kelas ke dalam dunia maya. Program Pembelajaran Daring Indonesia Terbuka dan Terpadu (PDITT) merupakan upaya nyata pemanfaatan teknologi dan komunikasi untuk memindahkan proses pembelajaran dari ruangkelas ke dunia maya. Mulai tahun 2014 Direktorat Jenderal Pendidikan Tinggi (Dikti) bekerja sama dengan beberapa Perguruan Tinggi menawarkan beberapa matakuliah melalui program PDITT. Layanan pembelajaran daring yang ditawarkan meliputi:

- 1) **Materi Terbuka:** sumber belajar yang utuh (aset digital), bebas diakses dan digunakan oleh siapapun, tanpa hak cipta (ketersediaan: daring, pemanfaatan: luring),
- 2) **Mata Kuliah Terbuka** mata kuliah utuh yang dapat diakses dan digunakan oleh siapa saja, sesuai standar proses dan standar isi, tanpa implikasi pengakuan kredit maupun sertifikasi oleh pengembang mata kuliah (ketersediaan: daring, pemanfaatan: luring),
- 3) **Mata Kuliah Daring:** mata kuliah utuh yang dapat diakses dan digunakan oleh siapa saja secara daring, sesuai standar proses dan standar isi, dan memperoleh sertifikat yang dapat dialihkreditkan. (ketersediaan: daring, pemanfaatan: daring dan luring).

Pembelajaran tatap muka dan pembelajaran daring merupakan cara pembelajaran untuk mencapai capaian pembelajaran yang telah ditentukan. Oleh karena itu, kualitas lulusan mahasiswa yang belajar secara tatap muka atau melalui daring harus sama. Untuk menilai hasil belajar mahasiswa khususnya dalam pembelajaran daring diperlukan alat ukur yang tepat dan sesuai dengan karakteristik pembelajaran daring.

Untuk menilai hasil belajar mahasiswa dalam pembelajaran daring diperlukan langkah-langkah sistematis agar asesmen yang dilakukan benar-benar dapat mengukur capaian belajar yang telah ditetapkan. Ada dua jenis asesmen yang perlu dikembangkan dalam pembelajaran daring yaitu asesmen mandiri dan asesmen oleh dosen. Asesmen mandiri adalah asesmen yang dilakukan sendiri oleh mahasiswa setelah mereka mempelajari satu atau lebih pokok bahasan, untuk mengukur capaian pembelajarannya. Asesmen mandiri dapat disediakan dalam bentuk kuiz atau latihan dan asesmen formatif. Asesmen mandiri disediakan agar mahasiswa dapat melakukan refleksi dan menilai kemampuan diri sendiri tentang hasil belajar yang telah dicapai. Asesmen oleh dosen adalah asesmen

yang dipersiapkan dan dilakukan oleh dosen untuk mengases hasil belajar mahasiswa. Asesmen yang dilakukan dosen untuk mengukur capaian hasil belajar mahasiswa dapat ditagih melalui tugas, ujian tengah semester, dan ujian akhir semester. Langkah-langkah yang harus dilakukan dosen untuk mempersiapkan asesmen hasil belajar mahasiswa meliputi: penentuan kompetensi yang akan diases, pembuatan perencanaan asesmen, pengembangan alat ukur berdasar kisi-kisi. Setelah kegiatan pembelajaran daring dilakukan maka diperlukan kegiatan evaluasi terhadap program pembelajaran daring tersebut. Hasil evaluasi pembelajaran daring berupa masukan-masukan untuk memperbaiki program pembelajaran.

Untuk membantu para dosen pengembang pembelajaran daring agar dapat merancang, melaksanakan, menilai hasil belajar mahasiswa dan mengevaluasi program pembelajaran daring diperlukan Panduan Pengukuran Kualitas Pelaksanaan Asesmen dan evaluasi Pembelajaran Daring.

Tujuan

Panduan ini dikembangkan dengan tujuan sebagai berikut:

- a. menjamin kualitas asesmen hasil belajar pada matakuliah Daring dilaksanakan dengan baik sesuai standar yang telah ditetapkan.
- b. mengevaluasi program pembelajaran daring.

Secara khusus, panduan ini bertujuan untuk:

1. Membantu dosen pengembang mata kuliah daring dalam:
 - 1) mempersiapkan asesmen mandiri yang meliputi pengembangan latihan (kuiz) dan asesmen formatif.
 - 2) Mempersiapkan dan melaksanakan asesmen yang meliputi pengembangan tugas, ujian tengah semester, dan ujian akhir semester.
 - 3) Mempersiapkan dan melaksanakan evaluasi program pembelajaran daring yang telah dilaksanakan.
2. Membantu mahasiswa:
 - 1) mendapat jaminan standar kualitas asesmen hasil belajar.
 - 2) mendapat jaminan layanan program pembelajaran daring yang berkualitas.

2. Pengertian dan Ruang Lingkup Asesmen dan Evaluasi Program Pembelajaran Daring

2.1. Pengertian asesmen dan evaluasi

Menurut Hanna (1993): *“Assessment is the process of collecting, interpreting, and synthesizing information to aid in decision making. Assessment synonymous with measurement plus observation. It concerns drawing inferences from these data sources.* Asesmen merupakan proses pengumpulan, interpretasi, dan sintesis informasi untuk mengambil keputusan. Pengambilan keputusan dalam asesmen dilakukan berdasar data yang diperoleh dari hasil pengukuran (berupa data kuantitatif) dipadu dengan data kualitatif yang diperoleh dari hasil pengamatan. Pengertian asesmen juga diberikan oleh Harlen (2006) yang menyatakan bahwa asesmen dalam konteks pendidikan merupakan proses pengumpulan bukti-bukti dan pembuatan pertimbangan untuk mengambil keputusan tentang hasil belajar siswa berkaitan dengan tujuan pembelajaran yang akan diukur. Tujuan utama asesmen lebih ditekankan pada upaya untuk meningkatkan pembelajaran yang dilakukan siswa (*student’s learning*) dan perkembangan belajar siswa daripada hanya sekedar menentukan hasil belajar siswa seperti dikatakan Morgan & O’Reilly (1999): *The primary purpose of assessment is to increase student’s learning and development rather than simply to grade or rank student performance.* Asesmen dalam pembelajaran dapat dilaksanakan dengan berbagai cara dan tujuan, tetapi menurut William & Leahy (2007) secara umum asesmen mempunyai tiga fungsi: 1) membantu meningkatkan pembelajaran, 2) menentukan hasil belajar siswa, dan 3) mengevaluasi kualitas program pembelajaran.

Berbagai jenis tagihan yang digunakan dalam asesmen pembelajaran antara lain: kuis, ulangan harian, tugas individu, tugas kelompok, ujian tengah semester, ujian akhir semester, laporan kerja dan lain sebagainya.

Jika kita bicara asesmen dan evaluasi dalam pembelajaran maka lingkup asesmen adalah pada individu siswa dalam kelas sedangkan lingkup evaluasi adalah seluruh komponen dalam program pembelajaran tersebut.

Evaluasi program pembelajaran merupakan evaluasi terhadap keseluruhan komponen program pembelajaran mulai perencanaan program pembelajaran termasuk kurikulum dan penilaian (asesmen) serta pelaksanaannya, pengadaan dan peningkatan kemampuan dosen, manajemen pendidikan, dan reformasi pendidikan secara keseluruhan. Evaluasi bertujuan untuk memberi

masuk kepada pengambil keputusan dalam rangka meningkatkan kualitas, kinerja, atau produktivitas suatu lembaga dalam melaksanakan programnya. Dalam hubungan dengan evaluasi program, Kaufman & Thomas (1980) mendefinikan evaluasi dengan menyatakan: *Evaluation is a process used to assess the quality of what is going on. Evaluation will provide quality control by determining the gap between what happened and what should have happened. Evaluation will tell us what is useful, what is not, and how to improve, what requires improvement.* Evaluasi merupakan suatu proses yang dimaksudkan untuk mengukur kualitas suatu program yang sedang berjalan. Evaluasi akan memberikan kontrol kualitas dengan cara menentukan adanya ketimpangan antara apa yang terjadi (*what happened*) dengan apa yang seharusnya terjadi (*what should have happened*). Evaluasi akan memberikan informasi tentang komponen mana yang dapat berfungsi dengan baik, mana yang tidak, dan bagaimana cara meningkatkan kinerja program tersebut, serta persyaratan apa yang diperlukan untuk melakukan perbaikan. Lebih lanjut Kaufman & Thomas (1980) menyatakan: ” *Evaluation is a process of helping to make things better than they are, of improving the situation*”. Evaluasi bertujuan untuk membantu memperbaiki program agar dapat berjalan lebih baik. Hal senada disampaikan oleh Madaus, Scriven, & Stufflebeam (1993) yang menyatakan: ...”*the most important purpose of evaluation is not to prove but to improve*”. Tujuan evaluasi tidak untuk mencari kesalahan dari program yang dievaluasi tetapi evaluasi dilakukan untuk memperbaiki kualitas program.

Banyak definisi evaluasi yang disampaikan oleh para ahli tetapi pada hakekatnya evaluasi selalu memuat masalah informasi dan kebijakan yaitu informasi tentang pelaksanaan dan keberhasilan suatu program yang selanjutnya digunakan untuk menentukan kebijakan berikutnya. Kalau Anda akan mengevaluasi program pembelajaran yang telah Anda lakukan maka Anda harus mengevaluasi pelaksanaan dan keberhasilan dari program pembelajaran yang telah Anda rencanakan. Hasil evaluasi pembelajaran diharapkan dapat mendorong dosen atau guru untuk mengajar lebih baik dan mendorong mahasiswa atau siswa untuk belajar lebih baik.

2.2. Asesmen dalam pembelajaran daring

Asesmen dalam pembelajaran daring merupakan proses penilaian yang dilakukan oleh dosen untuk menilai proses dan hasil belajar mahasiswa yang belajar melalui pembelajaran daring. Ada tiga kegiatan asesmen penting dilakukan oleh dosen selama pembelajaran daring.

1. Asesmen yang dilakukan oleh mahasiswa sendiri melalui asesmen mandiri (*self assessment*). Asesmen ini bertujuan memberi kesempatan kepada mahasiswa untuk

mengukur kemampuan diri sendiri dengan cara mengerjakan semua latihan dan asesmen formatif yang tersedia di bahan ajar. Jika ada konsep yang belum dipahami dengan baik mahasiswa dapat mempelajari kembali. Dalam hal ini dosen harus menyiapkan latihan, kuiz, dan asesmen formatif dalam pembelajarannya agar mahasiswa dapat melaksanakan asesmen mandiri. Latihan dan kuiz dimaksudkan untuk memantapkan konsep yang telah dipelajari mahasiswa. Instrumen latihan dapat berupa tes atau non-tes tergantung dari konsep yang akan dimantapkan. Sementara itu, asesmen formatif dimaksudkan agar mahasiswa dapat mengidentifikasi kesulitan belajar yang dialami dan dapat memperbaikinya kesulitan belajarnya berdasarkan hasil asesmen formatif yang dikerjakannya. Agar mahasiswa dapat mengidentifikasi kelemahan belajar dan dapat memperbaiki kelemahan belajar tersebut maka asesmen mandiri harus dilengkapi dengan kunci atau pedoman penskoran, petunjuk menghitung tingkat penguasaan kompetensi, dan petunjuk bagaimana cara memperbaiki kelemahan belajarnya. Instrumen asesmen formatif dibuat sesuai dengan kompetensi yang telah ditetapkan.

2. Asesmen yang dilakukan untuk mengukur proses dan hasil belajar mahasiswa. Asesmen ini dimaksudkan untuk mengukur capaian hasil belajar mahasiswa (asesmen sumatif). Asesmen jenis ini dapat dilakukan selama proses pembelajaran atau setelah proses pembelajaran selesai. Tagihan dalam asesmen sumatif dapat berupa tugas, unjuk kerja, portofolio, mid semester, dan Ujian Akhir semester. Jenis instrumen yang digunakan dalam asesmen sumatif disesuaikan dengan kompetensi yang akan diukur. Semua instrument yang digunakan dalam asesmen harus memenuhi syarat validitas dan reliabilitas dan dikonstruksi dengan baik.
3. Mengolah skor hasil asesmen untuk menentukan nilai akhir matakuliah. Nilai akhir matakuliah ditentukan dari komponen-komponen berikut: aktivitas mahasiswa selama mengikuti pembelajaran daring, nilai tugas, nilai ujian tengah semester, dan nilai ujian akhir semester. Pembobotan setiap komponen penentu nilai akhir matakuliah diatur sebagai berikut:

$$\text{NA} = 10\% \text{ nilai aktivitas mahasiswa} + 30\% \text{ Nilai Tugas} + 30\% \text{ Nilai Ujian Tengah Semester} + 40\% \text{ Nilai Ujian Akhir Semester.}$$

2.3. Prinsip asesmen pembelajaran daring

Agar asesmen yang Anda lakukan benar-benar dapat memberi gambaran yang sebenarnya tentang pencapaian hasil belajar mahasiswa maka dalam melakukan asesmen perlu diperhatikan prinsip-prinsip asesmen hasil belajar sebagai berikut.

Penilaian yang Anda lakukan harus berfungsi untuk mengukur ketercapaian siswa dalam pencapaian kompetensi seperti yang telah ditetapkan dalam kurikulum.

a. Valid

Penilaian yang Anda lakukan harus dapat mengukur apa yang seharusnya diukur. Untuk itu Anda memerlukan alat ukur yang dapat menghasilkan hasil pengukuran yang valid dan reliabel.

Contoh: pada akhir pembelajaran teknik mengelas mahasiswa diharapkan dapat mempraktekkan cara mengelas yang baik dan benar. Untuk mencapai kompetensi tersebut Anda tidak dapat menilainya hanya dengan menggunakan tes tertulis (*paper and pencil test*). Jika hanya itu yang Anda lakukan, Anda hanya akan dapat mengukur pengetahuan mahasiswa tentang mengelas. Agar Anda dapat mengetahui keterampilan mahasiswa dalam mengelas, Anda perlu menilai unjuk kerja mahasiswa. Untuk keperluan tersebut, Anda dapat memberi tugas (*task*) kepada mahasiswa untuk mempraktekkan cara mengelas yang baik dan benar. Untuk menilai keterampilan mahasiswa dalam mengelas, Anda harus membuat pedoman pengamatan yang dilengkapi dengan kriteria penskorannya (*rubric*). Kemudian gunakanlah rubrik tersebut untuk menilai kemampuan siswa dalam mengelas. Dengan cara seperti itulah kompetensi mahasiswa dalam mengelas dapat terukur dengan tepat.

b. Adil

Penilaian yang Anda lakukan harus adil untuk seluruh mahasiswa. Mahasiswa harus memperoleh kesempatan dan perlakuan yang sama. Contoh penilaian tidak adil yang sering kita temukan di lapangan, misalnya dalam tes tertulis guru menyediakan 10 butir soal. Semua mahasiswa diwajibkan mengerjakan butir soal nomor 1 – 5 dan diberi kebebasan untuk memilih 2 dari 5 butir soal nomor 6 – 10. Dari contoh tersebut tampak bahwa semua mahasiswa mendapat perlakuan yang sama hanya untuk mengerjakan butir soal nomor 1 – 5 tetapi tidak mendapat perlakuan yang sama untuk 2 butir soal pilihan yang diambil dari butir soal nomor 6 – 10. Hal ini akan menjadi masalah jika kompetensi yang diukur oleh butir soal nomor 6-10 adalah berbeda.

c. Objektif

Dalam menilai hasil belajar mahasiswa Anda harus dapat menjaga objektivitas proses dan hasil penilaian. Objektivitas penilaian dipengaruhi oleh unsur subjektivitas penilai. Unsur

subjektivitas dapat mempengaruhi penilaian pada saat pelaksanaan, penskoran, dan pengambilan keputusan hasil belajar siswa. *Hallo effect, carry over effect, order effect*, serta *mechanic and language effect* dapat menjadi penyebab tingginya unsur subjektivitas hasil penskoran.

d. Berkesinambungan

Penilaian yang Anda lakukan harus terencana, bertahap, teratur, terus menerus dan berkesinambungan untuk memperoleh informasi hasil belajar dan perkembangan belajar mahasiswa. Pengambilan keputusan pencapaian hasil belajar mahasiswa tidak boleh dilakukan hanya berdasar informasi hasil belajar mahasiswa pada tes akhir semester saja tetapi harus diputuskan berdasar informasi hasil belajar mahasiswa dari berbagai sumber yang diperoleh secara berkesinambungan. Hasil belajar harus dianalisis dan ditindaklanjuti dengan pemberian umpan balik sehingga dapat diperoleh catatan tentang perkembangan belajar mahasiswa. Informasi tersebut juga harus dapat dimanfaatkan untuk perbaikan pembelajaran pada semester berikutnya. Dengan demikian penilaian harus merupakan bagian integral dari pembelajaran. Dengan melakukan penilaian secara berkelanjutan, Anda tidak hanya melakukan penilaian dalam arti asesmen tetapi Anda juga dapat melakukan evaluasi terhadap program pembelajaran yang telah Anda laksanakan.

e. Menyeluruh

Prinsip menyeluruh dalam penilaian mengandung arti bahwa penilaian yang Anda lakukan harus mampu menilai keseluruhan kompetensi yang telah ditetapkan dalam Rancangan Acara Tutorial (RAT) maupun Satuan aktivitas Tutorial (SAT) yang mungkin meliputi ranah kognitif, afektif, dan psikomotor. Idealnya penilaian hasil belajar mahasiswa harus mampu mengukur keseluruhan kompetensi yang telah ditetapkan tetapi karena keterbatasan waktu yang disediakan untuk penilaian biasanya kita hanya mampu menilai ketercapaian kompetensi tertentu. Jika hal tersebut harus terjadi maka dosen harus dapat memilih kompetensi terpenting yang harus diukur.

f. Terbuka

Kriteria penilaian harus terbuka bagi berbagai kalangan sehingga keputusan hasil belajar mahasiswa jelas bagi pihak-pihak yang berkepentingan. Kriteria penilaian harus jelas, jika ada pihak-pihak yang mempertanyakan hasil penilaian maka evaluator harus mampu mempertanggungjawabkan hasil penilaiannya.

g. Bermakna

Hasil penilaian hendaknya bermakna bagi mahasiswa dan juga pihak-pihak yang berkepentingan. Hasil penilaian hendaknya dapat memberikan gambaran mengenai tingkat pencapaian hasil belajar mahasiswa, keunggulan dan kelemahan mahasiswa, minat, serta potensi mahasiswa dalam mencapai kompetensi yang telah ditetapkan.

2.4. Evaluasi program pembelajaran daring

Evaluasi program pembelajaran daring merupakan kegiatan pengumpulan data/informasi, analisis, dan pengambilan keputusan terhadap program pembelajaran daring. Evaluasi program pembelajaran daring dilakukan terhadap komponen-komponen pembelajaran daring seperti program aplikasi pembelajaran daring, mahasiswa, dosen, bahan sajian, layanan bantuan belajar (tutorial), pengelolaan program pembelajaran daring, kepuasan mahasiswa terhadap layanan pembelajaran daring, dan hasil belajar mahasiswa. Hasil evaluasi berupa masukan-masukan yang dapat dimanfaatkan oleh dosen atau pengambil kebijakan untuk menentukan keberlanjutan program.

Program pembelajaran daring dapat dievaluasi dengan menggunakan beberapa model evaluasi yang telah ditemukan oleh beberapa ahli atau menggunakan gabungan dari beberapa model evaluasi. Model-model yang dapat dimanfaatkan antara lain model evaluasi formatif-sumatif dari Scriven, model evaluasi CIPP dari Stufflebeam, atau model evaluasi pelatihan dari Kirkpatrick.

a. Model Evaluasi Formatif-Sumatif Scriven

Evaluasi formatif merupakan proses asesmen yang dilakukan terhadap program yang sedang berjalan untuk mengetahui kemajuan program dalam mencapai tujuan. Jika kemajuan program berjalan dengan baik dan mengarah pada pencapaian tujuan maka tidak perlu ada perbaikan tetapi jika ada komponen-komponen dari program yang tidak dapat berjalan seperti yang diharapkan maka pengembang program dapat memberikan umpan balik untuk memperbaiki komponen tersebut agar komponen program tersebut dapat berjalan seperti yang diharapkan. Menurut Scriven (Kaufman & Thomas, 1980) evaluasi formatif fokus pada empat pertanyaan berikut.

- 1) Apakah program berjalan sesuai dengan perencanaan ?
- 2) Apakah semua komponen program berfungsi efektif atau ada beberapa yang memerlukan perbaikan ?

- 3) Berdasarkan kemajuan data yang ada, haruskah program diterapkan ditempat lain ?
- 4) Apakah ada kejadian-kejadian penting yang harus digabungkan dalam struktur program ?

Dengan demikian evaluasi formatif sangat bermanfaat khususnya bagi pengembang program (*program designer*). Dengan melakukan evaluasi formatif maka jika terdapat komponen-komponen program yang tidak berjalan seperti yang diharapkan, pengembang program dapat memperbaiki, memodifikasi, atau membuat penyesuaian sehingga program dapat kembali berjalan seperti yang diharapkan.

Sebaliknya evaluasi sumatif merupakan evaluasi akhir program yang sangat bermanfaat untuk mengetahui sejauh mana tujuan program dapat dicapai. Menurut Scriven (Kauffman & Thomas, 1980) evaluasi sumatif dimaksudkan untuk menjawab pertanyaan-pertanyaan sebagai berikut.

- 1) Seberapa besar peserta program dapat memahami tentang materi yang diajarkan ?
- 2) Seberapa besar telah terjadi perubahan sikap dan atau perilaku pada peserta program setelah mengikuti program ini ?
- 3) Sejauh mana tujuan yang telah ditetapkan oleh pengembang program dapat tercapai ?
- 4) Apakah program benar-benar dapat membuat perbedaan pada peserta program sebelum dan sesudah mengikuti program ?
- 5) Apakah program dapat mencapai dan memenuhi kebutuhan yang telah teridentifikasi (*identified needs*) dan tujuan program ?
- 6) Apakah kesenjangan yang ada sebagai hasil *needs assessment* sudah dapat terisi ?
- 7) Apakah kinerja peserta program sudah berada pada *level* seperti yang ditetapkan dalam tujuan program ?

Pada saat melakukan evaluasi sumatif, perhatian utama harus ditekankan pada pengukuran terhadap hasil/efek utama dari program tersebut daripada mengukur pengaruh-pengaruh luar yang mungkin telah mempengaruhi peserta program. Untuk membedakan antara evaluasi formatif dan sumatif, Bob Stake (Madaus & Kellaghan, 2002) memberikan perumpamaan yang menarik dengan menyatakan: “*When the cook tastes the soup, that’s formative; when the guests taste the soup, that’s summative.*”

b. Model Evaluasi CIPP Stufflebeam

Model evaluasi CIPP (*Context, Input, Process and Product*) merupakan model evaluasi yang menekankan pada evaluasi: 1) konteks, 2) *input*, 3) proses, dan 4) produk. Model evaluasi CIPP dimaksudkan untuk memperbaiki program bukan untuk membuktikan, seperti dikatakan Stufflebeam (2002): “*the CIPP approach is based on the view that the most important purpose of*

evaluation is not to prove but to improve". Model evaluasi CIPP dikembangkan untuk membantu para manager dan para pelaksana program untuk mengambil empat macam keputusan yaitu:

- 1) Evaluasi konteks untuk membantu pengambilan keputusan tentang perencanaan (*planning*). Hasil evaluasi konteks sangat diperlukan untuk membuat perencanaan. Dengan adanya evaluasi konteks, perencana program dapat mengetahui kebutuhan (*need*) yang harus diperhatikan dalam membuat perencanaan serta dapat digunakan sebagai bahan pertimbangan untuk merumuskan tujuan program. Stufflebeam & Shinkfield (1985) menjelaskan bahwa evaluasi konteks dimaksudkan untuk mendefinisikan konteks penyelenggara program (*institutional context*) untuk mengetahui target, pengukuran *need*, mendiagnosis problem yang berhubungan dengan *need*, dan memberi usulan yang tepat untuk mengukur *need*. Evaluasi konteks dapat dilakukan dengan analisis sistem, survei, telaah dokumen, dengar pendapat, wawancara, dan tes diagnostik.
- 2) Evaluasi *input* untuk membantu pengambilan keputusan tentang struktur program yang akan dilaksanakan. Dari hasil evaluasi *input*, perancang program dapat menentukan sumber atau fasilitas yang diperlukan, merancang strategi pelaksanaan program, dan membuat perencanaan yang paling tepat untuk mencapai tujuan. Stufflebeam & Shinkfield (1985) menjelaskan bahwa evaluasi *input* dimaksudkan untuk mengidentifikasi dan mengukur kemampuan program, menyusun strategi, mendisain prosedur pelaksanaan program, menentukan anggaran, dan jadwal pelaksanaan program. Evaluasi *input* dapat dilakukan dengan cara mengumpulkan dan menganalisis sumber daya manusia dan sumber materi yang akan digunakan, menggunakan strategi pemecahan masalah, atau melakukan studi kelayakan.
- 3) Evaluasi proses untuk membantu pengambilan keputusan tentang pelaksanaan program. Hasil evaluasi proses dapat digunakan untuk mengetahui adanya hambatan pelaksanaan program serta upaya untuk memperbaiki program. Dengan melakukan evaluasi proses maka prosedur pelaksanaan program dapat dikontrol, dimonitor, dan diperbaiki. Menurut Worthen & Sanders (1987) evaluasi proses bertujuan untuk mengidentifikasi atau memprediksi proses, ketidaknormalan rancangan disain, atau ketidaknormalan pelaksanaan program. Evaluasi proses dapat dilakukan melalui kegiatan monitoring pelaksanaan program dengan cara melakukan pengamatan dan wawancara. Stufflebeam (Suharsimi Arikunto & Cepi Safruddin Abdul Jabar, 2009) mengusulkan beberapa pertanyaan pada saat melakukan evaluasi proses:
 - a) Apakah pelaksanaan program sesuai dengan jadwal ?
 - b) Apakah staf yang terlibat di dalam pelaksanaan program akan sanggup menangani kegiatan selama program berlangsung dan kemungkinan jika program dilanjutkan ?

- c) Apakah sarana dan prasarana yang disediakan dimanfaatkan secara maksimal ?
 - d) Hambatan-hambatan apa saja yang dijumpai selama pelaksanaan program dan kemungkinan jika program dilanjutkan ?
- 4) Evaluasi produk untuk membantu pengambilan keputusan tentang pencapaian tujuan program. Fitzpatrick, Sanders, & Worthen (2004) menyatakan evaluasi produk bertujuan untuk mengumpulkan gambaran dan pertimbangan tentang hasil program dan menghubungkannya dengan tujuan program dan informasi tentang konteks, *input*, dan proses, serta menilai ketercapaian tujuan program.

c. Model Evaluasi Kirkpatrick

Dalam mengevaluasi suatu program pelatihan, Kirkpatrick & Kirkpatrick (2006) menekankan pada empat *level* evaluasi yaitu:

- 1) evaluasi *level 1* yaitu evaluasi reaksi peserta terhadap pelaksanaan program pelatihan (*reaction level*).
- 2) evaluasi *level 2* yaitu evaluasi terhadap peserta setelah mengikuti program pelatihan (*learning level*)
- 3) evaluasi *level 3* yaitu evaluasi terhadap perilaku peserta setelah kembali ke tempat kerja (*behavior level*).
- 4) evaluasi *level 4* yaitu evaluasi terhadap hasil yang diperoleh institusi atau perusahaan setelah peserta yang dikirim untuk mengikuti pelatihan kembali ke tempat kerja (*result level*).

Pada umumnya evaluasi pelatihan yang banyak dilakukan adalah evaluasi *level* pertama yaitu evaluasi kepuasan peserta terhadap pelaksanaan program dan evaluasi *level* kedua yaitu evaluasi terhadap hasil pembelajaran. Kedua jenis evaluasi ini akan dilakukan setelah peserta selesai mengikuti seluruh rangkaian program pelatihan.

Evaluasi pada *level* reaksi dilakukan dengan cara mengevaluasi reaksi peserta terhadap pelaksanaan program pelatihan, seberapa besar kepuasan peserta terhadap pelaksanaan pelatihan yang telah diikutinya. Hasil evaluasi pada *level* reaksi akan digunakan untuk mengetahui efektivitas program dan untuk memperbaiki pelaksanaan program pelatihan pada masa yang akan datang. Untuk mengevaluasi kepuasan pelanggan, peserta diminta untuk memberikan reaksi terhadap pelaksanaan program pelatihan yang dapat diberikan dalam bentuk pendapat, komentar, sikap terhadap instruktur (penguasaan materi, penampilan, penggunaan metode/media, kemampuan mengelola pelatihan), fasilitas pelatihan (kualitas dan kelengkapan fasilitas pelatihan), dan pengelolaan kegiatan pelatihan (jadwal, makalah, pengaturan kegiatan pelatihan, pelayanan kepada peserta).

Evaluasi pada level kedua adalah evaluasi pembelajaran yang dicapai peserta setelah mengikuti pelatihan yang menyangkut penguasaan materi pelatihan (*what knowledge was learned*), keterampilan yang diperoleh (*what skills were developed or improved*), dan perubahan sikap peserta setelah mengikuti pelatihan (*what attitudes were changed*). Penguasaan materi pelatihan peserta program dapat diukur dengan menggunakan tes, perubahan sikap dapat diukur dengan menggunakan skala sikap, dan keterampilan peserta dapat diukur dengan menggunakan *performance test*.

Evaluasi tingkah laku yang merupakan evaluasi level tiga dalam model Kirkpatrick tidak dilakukan pada akhir program pelatihan tetapi dilakukan setelah peserta kembali ditempat kerja. Tujuan evaluasi level ketiga ini adalah untuk mengetahui apakah tambahan pengetahuan, keterampilan, dan perubahan sikap yang diperoleh selama pelatihan dapat diterapkan di tempat mereka bekerja. Dengan melakukan evaluasi pada level tiga ini akan dapat diketahui relevansi program pelatihan dengan kebutuhan nyata di tempat kerja.

Evaluasi hasil yang merupakan evaluasi level empat dari model Kirkpatrick difokuskan pada hasil akhir (*final result*) yang terjadi setelah peserta program selesai mengikuti suatu program dan kembali ke tempat kerja. Penekanan dari evaluasi hasil model Kirkpatrick adalah apakah setelah peserta program selesai mengikuti program pelatihan dan kembali ke tempat kerja dapat menerapkan hasil pelatihan sehingga membawa dampak positif terhadap perusahaan. Yang termasuk dalam kategori hasil akhir dari suatu program pelatihann adalah apakah setelah peserta *training* kembali ke tempat kerja berdampak positif terhadap perusahaan, misalnya adanya kenaikan produksi, peningkatan kualitas, penurunan biaya, penurunan kuantitas terjadinya kecelakaan kerja, penurunan *turnover* dan kenaikan keuntungan. Beberapa program mempunyai tujuan meningkatkan moral kerja maupun membangun *teamwork* yang lebih baik. Dengan kata lain, evaluasi dilakukan terhadap *impact* program. Tidak semua *impact* sebuah program dapat diukur dan juga membutuhkan waktu yang cukup lama. Oleh karena itu, evaluasi level empat ini lebih sulit dibandingkan dengan evaluasi pada level-level sebelumnya.

Untuk mengevaluasi program pembelajaran daring dapat dilakukan dengan menggunakan paduan model evaluasi CIPP Stufflebeam dan model evaluasi pelatihan Kirkpatrick. Komponen program yang dievaluasi meliputi: perencanaan program, pelaksanaan program, hasil belajar mahasiswa, dan kepuasan mahasiswa terhadap pelaksanaan program.

2.5. Ruang lingkup asesmen dan evaluasi pembelajarandaring

- 1) Ruang lingkup asesmen dalam pembelajaran daring meliputi:
 - a. Penyiapan instrument asesmen mandiri yang berupa latihan atau pemberian kuiz serta pengembangan asesmen formatif yang memenuhi syarat sebagai asesmen yang baik.
 - b. Penyiapan instrument untuk mengukur hasil belajar mahasiswa yang dapat berupa instrument untuk: tugas, unjuk kerja, portofolio, ujian tengah semester, dan atau ujian akhir semester.
 - c. Pelaksanaan asesmen
 - d. Hasil asesmen
- 2) Ruang lingkup evaluasi pembelajaran daring meliputi:
 - a. pengembangan disain untuk mengevaluasi komponen: 1) perencanaan program, 2) pelaksanaan program, 3) hasil program, 4) pengelolaan program pembelajaran daring, dan 5) kepuasan mahasiswa terhadap program pembelajaran daring.
 - b. pengembangan instrument untuk mengevaluasi kelimakomponen program tersebut.
 - c. Pelaksanaan evaluasi
 - d. Pengolahan dan analisis data
 - e. Pelaporan hasil evaluasi

2.6. Mutu dan rancangan asesmen daring

Berikut ini disajikan standar mutu dan instrumen untuk mengetahui kualitas asesmen dan evaluasi pembelajaran daring.

1) Standar Mutu untuk

Asesmen Mandiri	
a. Latihan atau Kuiz	<ol style="list-style-type: none">a. Instrumen harus dikonstruksi dengan baikb. Instrumen harus dapat digunakan untuk menguatkan konsep yang telah dipelajari mahasiswa.
2. Asesmen formatif	<ol style="list-style-type: none">a. Instrumen harus dikonstruksi dengan baikb. Dapat digunakan untuk mengidentifikasi kesulitan belajar yang dialami mahasiswa.c. Dilengkapi dengan kunci atau pedoman penskorand. Dilengkapi dengan cara menghitung tingkat penguasaan materi

Asesmen Mandiri	
	<p>mahasiswa.</p> <p>c. Dilengkapi dengan petunjuk untuk memperbaiki kelemahan belajar mahasiswa.</p>
Asesmen Hasil belajar Mahasiswa	
Tugas, Unjuk Kerja, Portofolio, Mid Test, UAS	<ol style="list-style-type: none"> 1. Dapat mengukur kompetensi yang telah ditetapkan dalam Silabus dan Satuan Acara Perkuliahan (SAP). 2. Dikembangkan berdasar kisi-kisi 3. Dikonstruksi dengan baik 4. Dilaksanakan (diadministrasikan) dengan baik 5. Pemeriksaan hasil asesmen dilakukan dengan memperhatikan persyaratan pemeriksaan hasil asesmen yang baik 6. Dapat dikelola dengan baik sesuai dengan karakteristik pembelajaran daring.
Evaluasi Program Pembelajaran Daring	
Perencanaan	<ol style="list-style-type: none"> 1. Evaluasi komponen perencanaan pembelajaran daring dilakukan terhadap aspek-aspek berikut: kualitas dan kemudahan penggunaan aplikasi pembelajaran daring, penyiapan materi pembelajaran, perencanaan tutorial, dan perencanaan asesmen hasil belajar. 2. Evaluasi setiap aspek perencanaan program pembelajaran daring dilakukan dengan menggunakan instrument yang valid dan reliabel
Pelaksanaan proses pembelajaran daring	<ol style="list-style-type: none"> 1. Proses pembelajaran berjalan sesuai dengan perencanaan. 2. Proses pembelajaran berlangsung dengan pendekatan student active learning. 3. Proses pembelajaran dievaluasi dengan menggunakan instrument yang tepat.
Pengelolaan pembelajaran daring	<ol style="list-style-type: none"> 1. Adanya sosialisasi pembelajaran daring ke mahasiswa. 2. Mahasiswa mengetahui tatacara dan aturan main dalam pembelajaran daring. 3. Pengelolaan pembelajaran daring dievaluasi dengan cara yang tepat.

Asesmen Mandiri	
Hasil Program	<ol style="list-style-type: none"> 1. Hasil program berupa hasil belajar mahasiswa 2. Hasil belajar mahasiswa berupa penguasaan kompetensi yang telah ditentukan dalam silabus dan SAP. 3. Pengukuran kompetensi hasil belajar mahasiswa dilakukan dengan menggunakan instrument yang valid dan reliabel.
Kepuasan Mahasiswa	<ol style="list-style-type: none"> 1. Kepuasan mahasiswa terhadap layanan dalam pembelajaran daring merupakan salah satu tolok ukur keberhasilan program. Kepuasan mahasiswa terhadap layanan program pembelajaran daring meliputi: kepuasan mahasiswa terhadap persiapan pelaksanaan program, pelaksanaan program, dan hasil program. 2. Kepuasan mahasiswa terhadap layanan program pembelajaran daring dievaluasi dengan menggunakan instrument yang valid dan reliabel.

2) Pengembangan Instrumen Asesmen

Berikut ini disajikan daftar cek untuk menilai kualitas asesmen mandiri, asesmen hasil belajar dan evaluasi program pembelajaran daring.

Beri tanda cek (V) pada kolom Ya atau Tidak sesuai dengan hasil penilaian Anda terhadap pernyataan yang ada.

Aspek yang dievaluasi	Pernyataan	Ya	Tidak
Asesmen Mandiri			
1. Latihan dan atau Kuiz	<ol style="list-style-type: none"> 1. Latihan dan atau kuiz dikonstruksi dengan jelas dan tegas sehingga tidak ada salah tafsir. 2. Latihan dan atau kuiz dapat digunakan untuk menguatkan konsep yang telah dipelajari mahasiswa. 		
2. Asesmen formatif	<ol style="list-style-type: none"> 1. Asesmen formatif dikonstruksi sesuai dengan aturan tata tulis instrument yang baik. 2. Asesmen formatif dapat digunakan untuk mengidentifikasi konsep yang belum dipahami mahasiswa. 		

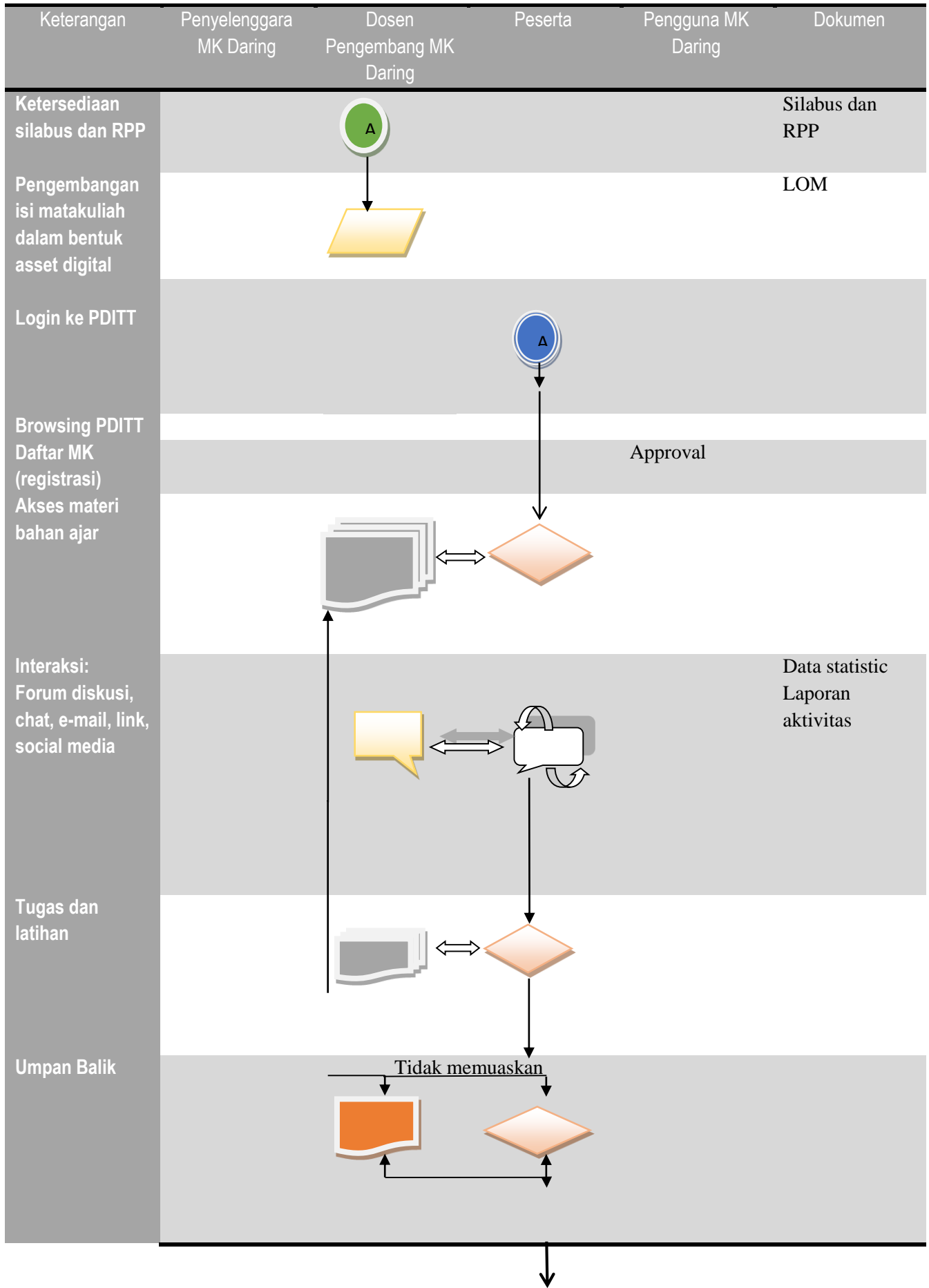
Aspek yang dievaluasi	Pernyataan	Ya	Tidak
	<p>3. Pilih yang sesuai:</p> <p>a. Jika asesmen formatif dalam bentuk tes pilihan ganda maka telah dilengkapi dengan kunci jawaban.</p> <p>b. Jika asesmen formatif dalam bentuk tes uraian maka telah dilengkapi dengan pedoman penskoran.</p> <p>c. Jika asesmen formatif dalam unjuk kerja atau portofolio maka telah dilengkapi dengan rubric.</p> <p>4. Asesmen formatif dilengkapi dengan cara menghitung tingkat penguasaan materi mahasiswa.</p> <p>5. Asesmen formatif dilengkapi dengan petunjuk untuk memperbaiki kelemahan belajar mahasiswa.</p>		
Asesmen Hasil Belajar Mahasiswa			
Tugas, Unjuk Kerja, Portofolio, Mid Test, UAS	<p>1. Instrumen yang digunakan dalam asesmen sumatif dapat digunakan untuk mengukur kompetensi yang telah ditetapkan dalam silabus dan SAP</p> <p>2. Dosen telah mengembangkan kisi-kisi asesmen sumatif.</p> <p>3. Instrumen asesmen sumatif dikembangkan berdasar kisi-kisi.</p> <p>4. Instrumen asesmen sumatif telah dikonstruksi berdasar aturan tata tulis pengembangan instrument yang baik.</p> <p>5. Asesmen sumatif dilaksanakan dengan pengawasan yang ketat.</p> <p>6. Telah diyakinkan bahwa mahasiswa peserta asesmen sumatif adalah mahasiswa yang benar (bukan joki).</p> <p>7. Ruang pelaksanaan asesmen sumatif telah diatur sedemikian rupa sehingga tidak memungkinkan mahasiswa melakukan kecurangan dalam mengikuti</p>		

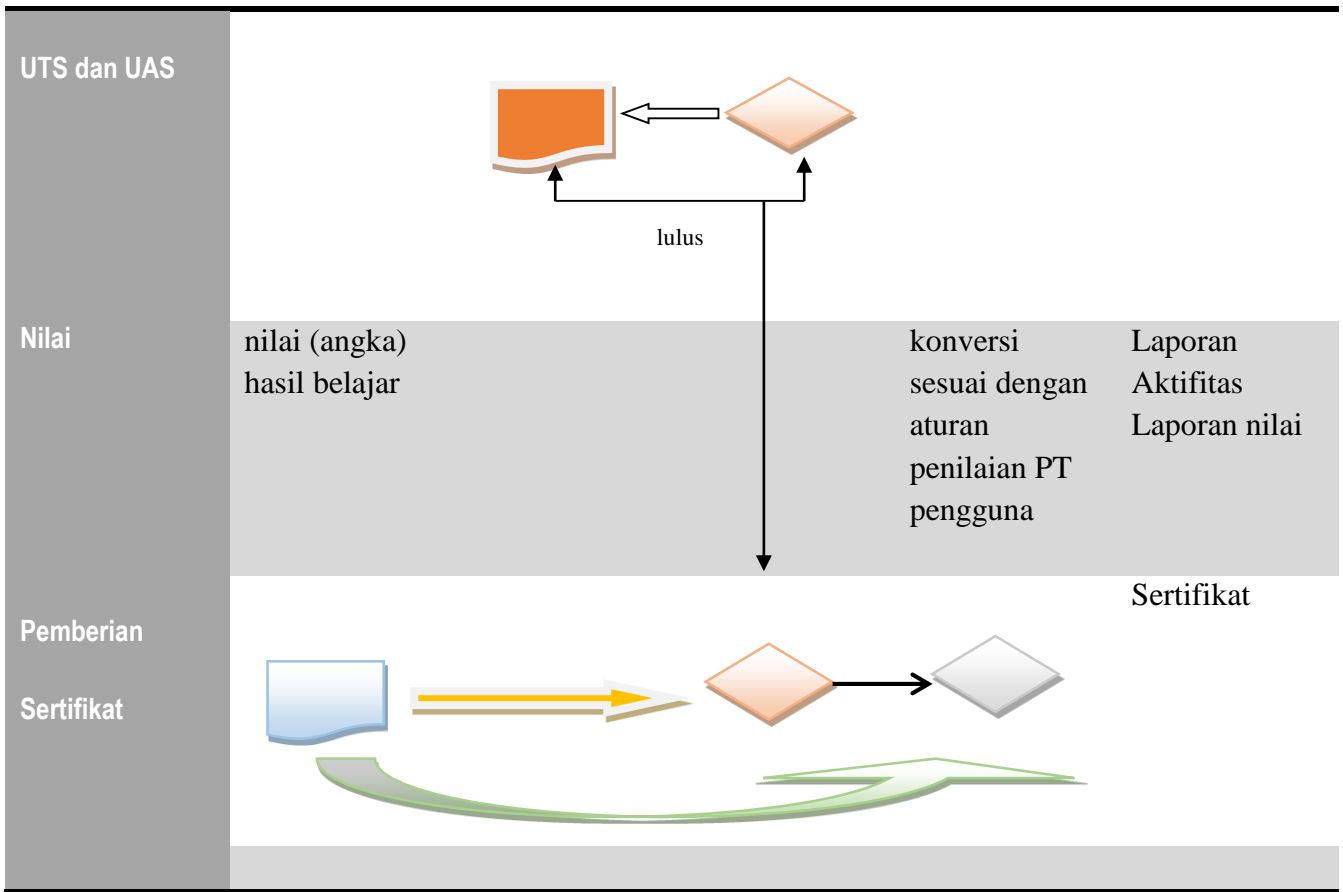
Aspek yang dievaluasi	Pernyataan	Ya	Tidak
	<p>ujian.</p> <p>8. Pemeriksaan hasil asesmen dilakukan secara objektif.</p> <p>9. Asesmen sumatif yang dilakukan dapat dikelola dengan baik sesuai dengan karakteristik pembelajaran daring.</p>		
Evaluasi Program Pembelajaran daring			
Perencanaan	<ol style="list-style-type: none"> 1. Aplikasi pembelajaran daring dapat digunakan dengan mudah oleh mahasiswa. 2. Aplikasi pembelajaran daring dapat digunakan untuk mengelola semua kegiatan pembelajaran daring. 3. Aplikasi pembelajaran daring cukup handal (tidak mudah hank) pada saat digunakan selama pembelajaran daring berlangsung. 4. Materi pembelajaran daring telah disiapkan lengkap sebelum pembelajaran daring dimulai. 5. Dosen telah menyusun Rancangan Aktivitas Tutorial (RAT). 6. Dosen telah menyusun Satuan Acara Tutorial (SAT). 7. Dosen telah mengembangkan kisi-kisi asesmen sumatif. 8. Instrumen asesmen sumatif dikembangkan berdasar kisi-kisi. 9. Instrumen asesmen sumatif telah dikonstruksi sesuai dengan aturan tata tulis pengembangan instrument yang baik. 		

Aspek yang dievaluasi	Pernyataan	Ya	Tidak
Pengelolaan pembelajaran daring	<ol style="list-style-type: none"> 1. Sosialisasi adanya pembelajaran daring ke mahasiswa telah dilakukan. 2. Mahasiswa mengetahui tatacara dan aturan main dalam mengikuti pembelajaran daring. 3. Mahasiswa menerima jadwal pelaksanaan tutorial pembelajaran daring sebelum tutorial dimulai. 4. Mahasiswa memperoleh layanan pembelajaran daring secara lengkap. 		
Hasil Program	<ol style="list-style-type: none"> 1. Instrumen yang digunakan dalam asesmen sumatif mempunyai validitas dan reliabilitas yang tinggi. 2. Setelah asesmen sumatif mahasiswa dapat menguasai kompetensi yang ditetapkan dalam silabus dan SAP. 		
Kepuasan Mahasiswa	<ol style="list-style-type: none"> 1. Mahasiswa puas terhadap persiapan pembelajaran daring yang dilakukan pengelola. 2. Mahasiswa puas terhadap layananpengelola pembelajaran daring. 3. Mahasiswa puas terhadapkualitas pembelajaran daring yang dilakukan dosen. 4. Mahasiswa puas terhadap pelaksanaan pembelajaran daring. 5. Mahasiswa puas terhadap hasil belajar yang diperoleh dalam pembelajaran daring. 		

Catatan: Banyaknya tanca cek pada kolom Ya menunjukkan bahwa kualitas asesmen dan evaluasi pembelajaran daring semakin baik.

3. Standar Operating Procedure





Daftar Pustaka

- Fitzpatrick, J.L., Sanders, J.R. & Worthen, B.R. (2004). *Program Evaluation: Alternative approaches and practical guidelines*(3rd ed.). Boston: Pearson.
- Hanna, G.S. (1993). *Better teaching trough better measurement*. New York: Harcourt Brace Jovanovich College Pub.
- Harlen, W. (2006). On the relationship between assessment for formative and summative purposes. Dalam Gardner (Eds). *Assessment and Learning*. London: Sage Publications.
- Kaufman, R. & Thomas, S. (1980). *Evaluation without fear*. New York: New Viewpoints.
- Kirkpatrick, D.L. & Kirkpatrick, J.D. (2006). *Evaluating Training Programs: the four levels* 3rd ed. San Francisco: Berret-Koehler Publishers, Inc.
- Madaus, G.F. & Kellaghan, T. (2002). Models, metaphors, and definitions in evaluation. In Stufflebeam, D.L., Madaus, & Kellaghan, T (Ed.).. *Evaluation models: Viewpoints on educational and human services evaluation second edition*. New York: Kluwer academic Publishers.
- Madaus, G.F. & Scriven, M.S. & Stufflebeam, D.L. (1993). *Evaluation models: Viewpoints on educational and human services evaluation*. Boston: Kluwer-Nijhoff Publishing.
- Morgan, C. & O'Reilly, M. (1999). *Assessing open and distance learners*. London: Kogan Page Limited.
- Suharsimi Arikunto & Cepi Safruddin Abdul Jabar. (2009). *Evaluasi Program Pendidikan: Pedoman teoretis praktis bagi mahasiswa dan praktisi pendidikan. (Edisi kedua)*. Jakarta: Bumi Aksara.
- Stufflebeam, D.L. (2002). The CIPP model for evaluation. In Stufflebeam, D.L., Madaus, & Kellaghan, T (Ed.). *Evaluation models: Viewpoints on educational and human services evaluation second edition*. New York: Kluwer academic Publishers.
- Stufflebeam, D.L. & Shinkfield, A.J. (1985). *Systematic evaluation*. Boston: Kluwer Nijhof Publishing.
- Wiliam, D. & Leahy, S. (2007). A theoretical foundation for formative assessment. Dalam McMillan, J.H. (Ed.). *Formative Classroom Assessment: Theory into Practice*. New York: Teachers College Press.
- Worthen, B.R. & Sanders, J.R. (1987). *Educational Evaluation: Alternative approaches and practical guidelines*. New York: Longman

