

# VALIDITAS, RELIABILITAS, DAN KEPRAKTISAN UJIAN MELALUI OBSERVASI DAN BENTUK LISAN BAGI KELAS BESAR PADA PRODI PAI FAKULTAS TARBIYAH DAN BAHASA IAIN SURAKARTA

*Suparmin, Usman Abu Bakar, Giyoto, Ahmad Fauzi\**

## **Abstrak:**

*Penelitian ini merupakan penelitian deskriptif terhadap pelaksanaan ujian mandiri (non-terjadual) dosen pengampu matakuliah paralel (kelas besar/ lebih dari 30 mahasiswa dalam satu kelas) pada Prodi PAI semester genap tahun 2011/2012. Penelitian ini bertujuan untuk mendeskripsikan validitas, reliabilitas dan kepraktisan soal dan penyelenggaraannya. Data diperoleh dari angket dan observasi terhadap mahasiswa. Hasil penelitian menunjukkan bahwa: a) Tingkat validitas soal ujian lisan dilihat dari butir-butir pertanyaannya cukup mengukur kompetensi mata kuliah yang direncanakan, namun seringkali didapati soal yang sama pada beberapa mahasiswa untuk kelas paralel. Demikian pula dalam pelaksanaan ujian kecenderungan soal berkembang sesuai situasi dan kondisi Dosen pada waktu itu yang mengakibatkan mahasiswa kesulitan untuk menjawab pertanyaan yang diberikan karena lebih sulit dari materi yang telah diajarkan atau tidak konsisten terhadap kompetensi yang dituju. Tes lisan dengan cara tidak terjadwal serempak dan mandiri sangat baik untuk menanyakan perbedaan-perbedaan pendapat (judgment) perorangan terhadap kasus tertentu. Tetapi tes lisan tidak efektif untuk menanyakan kompetensi referensial dan normatif sebagaimana dalam buku sumber. b) Reliabilitassoal yang diujikan rendah hal ini dipengaruhi oleh tingkat subyektifitas dosen sangat tinggi karena kondisi psikologisnya dan respon mahasiswa terhadap soal yang diujikan. c) Tingkat kepraktisan soal ujian lisan juga rendah, karena soal rawan bocor dan untuk menjawabnya memerlukan waktu yang lebih longgar dari jadwal akademik yang telah ditentukan.*

**Kata Kunci:** *Validitas soal, subyektifitas dosen, kelas paralel, ujian lisan*

---

\* Institut Agama Islam (IAIN) Surakarta.

## PENDAHULUAN

Proses pendidikan merupakan satuan interaksi yang saling menentukan, menunjang, dan saling berkorespondensi antara rencana pembelajaran, proses pembelajaran, dan pengukuran hasil capaian pembelajaran. Evaluasi dilakukan melalui pengukuran dan penilaian sebagai tolok ukur keberhasilan suatu pembelajaran pada kompetensi tertentu. Evaluasi juga dipakai untuk alat standarisasi input, proses, maupun output. Evaluasi diselenggarakan dengan teknik observasi, tes lisan, tes tulis, quis, serta angket.

Evaluasi yang dilakukan dengan teknik observasi dan tes lisan sangat rentan sekali dengan berbagai hal yang subjektif yang mengurangi tingkat validitas soal, reliabilitas alat evaluasi, dan kepraktisannya terutama untuk kelas besar dan terjadwal paralel.

Berdasarkan beberapa pertimbangan di atas sangat perlu dilakukan penelitian untuk melihat secara dekat dan faktual mengenai evaluasi melalui observasi dan lisan yang dilaksanakan pada program studi yang memiliki jadwal perkuliahan yang paralel (satu matakuliah lebih dari satu kelas/ kelas besar) dengan kata lain kelas yang jumlah mahasiswanya lebih dari 30 orang. Program studi Pendidikan Agama Islam pada Fakultas Tarbiyah dan Bahasa merupakan fakultas yang paling besar dilihat dari jumlah mahasiswa dan dosen yang mengajar, yakni 540 kelas dan sekitar 1300 SKS setiap minggunya dengan jumlah mahasiswa 3400 orang dengan dosen yang ditugaskan sekitar 113 orang baik dosen tetap maupun dosen luar biasa.

Dari identifikasi masalah di atas dapat dirumuskan masalahnya sebagai berikut: Sejauh mana tingkat Validitas, Reliabilitas, dan Kepraktisan ujian melalui observasi dan bentuk lisan bagi kelas besar dan terjadwal paralel pada Program Studi Pendidikan Agama Islam Fakultas Tarbiyah dan Bahasa Semester Genap Tahun Akademik 2011/2012?

Penelitian ini bertujuan untuk menemukan tingkat validitas, reliabilitas, dan kepraktisan soal ujian pada matakuliah yang terjadwal paralel yang pelaksanaan ujiannya diselenggarakan secara tidak terjadwal/mandiri oleh dosen pengampu di Program Studi Pendidikan Agama Islam Fakultas Tarbiyah dan Bahasa.

Pada penelitian ini, dilaksanakan pada program studi Pendidikan Agama Islam, selama tiga bulan Mei sampai Juli 2012 pada Semester Genap tahun akademik 2011/2012.

Sumber data penelitian ini adalah: 1. Dokumen soal dan kompetensi matakuliah yang diujikan secara lisan, 2. Informasi yang diperoleh dari mahasiswa yang diuji lisan; 3. Mahasiswa yang diteliti adalah mahasiswa yang telah menempuh matakuliah terjadual secara paralel dan diujikan secara lisan oleh dosen pengampu.

Pengumpulan data ini dilakukan dengan angket, dan dokumen. Kedua teknik ini dipakai untuk mendapatkan validitas, reliabilitas, dan kepraktisan data yang diperoleh. Catatan ini dibuat secara langsung dan diklasifikasikan, direduksi, dan dianalisis sejalan dengan berlangsungnya pengambilan data.

Bentuk dan cara pengklasifikasian, reduksi, dan analisis; dalam penelitian ini, dibuat berdasarkan pada ciri data yang diperoleh selama penelitian berlangsung di lapangan, melalui data angket berdasarkan data dokumen yang diperoleh.

Analisis data dilakukan berdasarkan data angket dari mahasiswa yang berkenaan dengan isi, cara, dan kerangka pelaksanaan ujian lisan atau ujian yang tidak terjadual secara serempak oleh fakultas. Analisis ini dikategorikan menjadi tiga hal, yakni analisis yang mendeskripsikan validitas, reliabilitas, dan kepraktisan ujian non terjadual. Dari setiap kategori di atas terdiri dari beberapa item pertanyaan yang dijawab oleh responden yang merupakan mahasiswa yang telah diuji non terjadual.

## PEMBAHASAN

### Evaluasi Pendidikan

Evaluasi berasal dari kata *evaluation* yang artinya suatu upaya untuk menentukan nilai atau jumlah. Kata-kata yang terkandung di dalam defenisi tersebut pun menunjukkan bahwa kegiatan evaluasi harus dilakukan secara hati-hati, bertanggung jawab, menggunakan strategi, dan dapat dipertanggung jawabkan. Evaluasi dilaksanakan untuk menyediakan informasi tentang baik atau buruknya proses dan hasil kegiatan tersebut.

Zainul dan Nasution menyebutkan bahwa<sup>1</sup> evaluasi merupakan proses pengambilan keputusan dengan menggunakan informasi yang diperoleh melalui hasil belajar, baik melalui instrumen tes maupun non tes. Sedangkan menurut Purwanto<sup>2</sup> evaluasi selalu menyangkut

<sup>1</sup>Zainul A., *Alternatif Assesment* (Jakarta: Dirjen Dikti,2001).

<sup>2</sup>Ngalim Purwanto, *Prinsip-Prinsip dan Teknik Evaluasi Pengajaran* (Bandung: Remaja Rosdakarya, 2010), 4.

pemeriksaan tujuan yang ditetapkan. Evaluasi bisa dilakukan atas hasil atau proses, evaluasi atas hasil dengan melihat pencapaian tujuan pada hasil kegiatan.

Berdasarkan pengertian di atas dapat disimpulkan bahwa evaluasi merupakan suatu kegiatan yang bertujuan untuk mengetahui dan memperoleh informasi mengenai perubahan dan perkembangan peserta didik dalam mengikuti kegiatan proses belajar mengajar. Jadi pengertian evaluasi dilihat dari sudut pembelajaran adalah suatu proses yang sistematis yang berusaha untuk mengetahui tingkat keberhasilan mahasiswa dan pencapaian tujuan pembelajaran yang telah ditetapkan dan sekaligus memberikan gambaran tentang keefektifan pembelajaran.

Evaluasi pendidikan dapat diartikan sebagai pengukuran atau penilaian hasil belajar-mengajar, padahal antara keduanya punya arti yang berbeda meskipun saling berhubungan. Mengukur adalah membandingkan sesuatu dan satu ukuran (kuantitatif), sedangkan menilai berarti mengambil satu keputusan terhadap sesuatu dengan ukuran baik-buruk (kualitatif). Evaluasi pendidikan merupakan proses/kegiatan untuk memperoleh informasi berupa umpan balik (*feedback*) bagi penyempurnaan pendidikan; menentukan kemajuan pendidikan, dibandingkan dengan tujuan yang telah ditentukan.

Evaluasi pendidikan memberikan manfaat baik siswa/peserta pendidikan, pengajar maupun manajemen. Dengan adanya evaluasi, peserta didik dapat mengetahui sejauh mana keberhasilan yang telah dicapai selama mengikuti pendidikan.

## 1. Prinsip Evaluasi Pendidikan

Dimiyati dan Mujiono menyebutkan bahwa<sup>3</sup> evaluasi yang akan dilakukan juga harus mengikuti prinsip *kesahihan* (*valid*), *keterandalan* (*reliabilitas*), dan *praktis*.

### a) Validitas (*kesahihan*)

Sebuah evaluasi dikatakan valid jika evaluasi tersebut secara tepat, benar, dan sah telah mengungkapkan atau mengukur apa yang seharusnya diukur. Agar diperoleh hasil evaluasi yang sah, dibutuhkan instrumen yang memiliki/memenuhi syarat kesahihan suatu instrumen evaluasi. Sahih atau tidaknya

---

<sup>3</sup>Dimiyati dan Mudjiono, *Belajar dan Pembelajaran* (Jakarta: Rineka Cipta, 2006), 194-199.

evaluasi tersebut ditentukan oleh faktor-faktor instrumen evaluasi itu sendiri, administrasi evaluasi dan penskoran, serta respon-respon siswa.<sup>4</sup> Kesahihan instrumen evaluasi diperoleh melalui hasil pemikiran dan pengalaman. Dari dua cara tersebut, diperoleh empat macam kesahihan yang terdiri atas kesahihan isi (*content validity*), kesahihan konstruksi (*contruction validity*), kesahihan ada sekarang (*concurrent validity*), dan kesahihan prediksi (*prediction validity*).<sup>5</sup>

b) *Reliabilitas*

Reliabilitas (keterandalan) evaluasi berhubungan dengan masalah kepercayaan yaitu tingkat kepercayaan bahwa suatu evaluasi mampu memberikan hasil yang tepat. Maksud dari pernyataan ini adalah jika suatu evaluasi dilakukan pada subjek yang sama evaluasi senantiasa menunjukkan hasil evaluasi yang sama atau sifatnya *ajeg* dan stabil.

Keterandalan dipengaruhi oleh beberapa faktor, yaitu:

- a) Panjang tes (*length of test*). Panjang tes berhubungan dengan banyaknya butir tes. Pada umumnya lebih banyak butir tes, lebih tinggi keterandalan evaluasi. Hal ini terjadi karena makin banyak soal tes, makin banyak sampel yang diukur.
- b) Sebaran skor (*spread of scores*). Besarnya sebaran skor akan membuat kemungkinan perkiraan keterandalan lebih tinggi menjadi kenyataan.
- c) Tingkat kesulitan tes (*difficulty of test*). Tes yang paling mudah atau paling sukar untuk anggota-anggota kelompok yang mengerjakan cenderung menghasilkan skor tes keterandalan yang lebih rendah. Hal ini disebabkan antara hasil tes yang mudah dan sulit keduanya dalam suatu sebaran skor yang terbatas.
- d) Objektivitas (*objektivity*). Objektivitas suatu tes menunjuk kepada tingkat skor kemampuan yang sama (yang dimiliki oleh para siswa) dan memperoleh hasil yang sama dalam mengerjakan tes.

---

<sup>4</sup>Gronlund dalam Dimiyati dan Mudjiono, *Belajar dan Pembelajaran* (Jakarta: Rineka Cipta, 2006), 195.

<sup>5</sup>Suharsimi Arikunto, dan Safruddin Abdul Jabar, *Evaluasi Program Pendidikan Pedoman Praktis Bagi Mahasiswa dan Praktisi Pendidikan* (Jakarta: Bumi Aksara, 2010), 64.

e) *Kepraktisan*

Kepraktisan suatu evaluasi bermakna bahwa kemudahan-kemudahan yang ada pada instrumen evaluasi baik dalam mempersiapkan, menggunakan, menginterpretasi, memperoleh hasil, maupun kemudahan dalam menyimpan.

Faktor-faktor yang mempengaruhi kepraktisan instrumen evaluasi meliputi: a). kemudahanmengadministrasi; b) waktu yang disediakan untuk melancarkan kegiatan evaluasi; c)kemudahan menskor; d)kemudahan interpretasi dan aplikasi; e)tersedianya bentuk instrumen evaluasi yang ekuivalen atau sebanding.

## TEMUAN PENELITIAN

### Validitas Soal yang Diujikan

Pada item pertama jawaban mahasiswa mengungkapkan bahwa dosen yang mengajar matakuliah yang diujikan secara tidak tertulis dan terjadwal telah menyampaikan dengan baik materi-materi perkuliahnya sesuai dengan kompetensi matakuliah masing-masing, yakni sesuai dengan rambu-rambu kompetensi mata kuliah yang diampu. Kompetensi matakuliah telah disampaikan di awal semester ketika dosen mulai mengajar dan menyusun kontrak belajar dengan mahasiswa.

Sejumlah empat belas mahasiswa menyampaikan bahwa mahasiswa menerima materi kuliah sebagaimana kompetensi dasar matakuliah yang diampu. Matakuliah disampaikan secara *up-date*, yang berarti bahwa konsep dan materi kuliah dihubungkan dengan isu-isu yang sedang terjadi saat matakuliah disampaikan sehingga mahasiswa mampu merekonstruksi pemahamannya dengan kuat dan lama tertanam. Mengenai kesesuaian soal tes dengan materi yang diajarkan hal ini tampak dari jawaban 12 mahasiswa bermakna bahwa soal ujian yang diujikan sudah sesuai dengan apa yang diajarkan dan sesuai dengan Standar Kompetensi yang ditentukan.

Lebih lanjut mahasiswa juga mengungkapkan bahwa soal yang diujikan lebih sulit dari materi yang diajarkan. Hal ini disebabkan beberapa hal: *pertama*, mahasiswa kesulitan dalam mengerjakan soal disebabkan kurang sungguh-sungguh dalam belajar; *kedua*, beberapa soal ujian menunjukkan tidak terkait dengan materi yang diajarkan.

Walaupun soal ujian ini bermaksud mengukur penalaran mahasiswa untuk mengetahui seberapa jauh pemahamannya tentang kondisi terkini. Kemungkinan soal yang diberikan berupa penalaran oleh mahasiswa sendiri dan jauh dari materi yang disampaikan sebagai upaya menggali kemampuan mahasiswa dalam mengkaitkan materi yang telah diajarkan dengan permasalahan sesuai kondisi kekinian. Untuk hal ini seharusnya dosen dalam menyampaikan materi pembelajaran dikaitkan kondisi terkini sehingga ketika dibuat soal ujian mahasiswa tidak mengalami kesulitan.

*Ketiga*, soal belum mewakili materi yang diajarkan. Keterwakilan soal yang diujikan belum mewakili materi yang selama ini diberikan. Sebab soal tidak menanyakan materi secara menyeluruh. Soal ujian hanya sekitar 5–10 pertanyaan yang disesuaikan dengan kompetensi yang ingin dicapai. Tentunya tidak memungkinkan jika soal menanyakan materi secara menyeluruh dan berjumlah banyak, sebab ketersediaan waktu ujian juga terbatas. Untuk ujian praktik pada mata kuliah psikologi jumlah soal yang diujikan dianggap cukup mewakili materi yang diajarkan dan soal ujian menjadi bekal untuk pembelajaran ketika mengajar di kelas bagi calon pendidik nantinya.

Beberapa catatan untuk ujian non tulis berupa tes secara berkelompok sekitar 4-5 orang, pada umumnya kelompok terakhir cenderung waktunya tidak mencukupi. Disebabkan jadwal ujian non tulis sama dengan ujian tulis, padahal kebutuhan waktunya berbeda. Untuk ujian lisan berupa tes secara kelompok jika jumlah mahasiswanya banyak maka waktunya cenderung tidak mencukupi atau kurang efektif evaluasi dengan teknik tes lisan secara kelompok.

### **Reliabilitas Soal yang Diujikan**

Tes lisan dengan cara tidak terjadwal serempak dan mandiri sangat baik ketika diujikan untuk menanyakan perbedaan-perbedaan pendapat atau *judgment* perorangan terhadap kasus tertentu. Tetapi ketika ujian tersebut menanyakan kompetensi referensial dan normatif sebagaimana dalam buku sumber, tes tersebut kurang efektif untuk dilakukan atau tidak dapat diberlangsungkan. Ada beberapa hal yang membawa tes tersebut tidak reliabel karena mahasiswa yang lebih awal ujian saling memberitahu soal-soal yang telah diteskan kepada mahasiswa yang belum mendapat tes, sehingga bobot kesulitan

soalnya tidak merata bagi semua kelas yang diuji. Ada perlakuan yang tidak sama antar mahasiswa yang mengikuti matakuliah tersebut. Tes jenis ini sangat rentan terhadap ketidakadilan perlakuan terhadap mahasiswa yang diujikan dalam kelas yang sama. Apalagi matakuliah tersebut diikuti oleh ratusan peserta. Sehingga untuk tes lisan ini diperlukan strategi yang sangat baik untuk memperlakukan mahasiswa dengan bobot kesulitan tes yang relatif sama sehingga terdapat keadilan dalam banyak hal/aspek evaluasi, seperti keadilan waktu, sikap penguji, keadilan bobot soal, keadilan parameter penilaian yang konsisten atau keadilan konsistensi parameter/indikator penilaian.

Tes lisan dimungkinkan sangat tinggi akan bersifat terbuka terhadap isu-isu atau *hello effect* dari luar materi ujian yang tidak mengukur kompetensi mahasiswa. Sehingga konsistensi tujuan tes bisa berubah di tengah jalan karena faktor lingkungan tempat tes diselenggarakan. Sehingga penguji dengan mudah merubah strategi ujian sesuai situasi saat ujian. Ujian ini akan menjadi lebih tidak reliabel ketika banyak kelas mata kuliah tertentu diampu oleh satu dosen, sehingga dosen tersebut dimungkinkan secara psikologis berubah-ubah sikap karena capek atau dipengaruhi sikap mahasiswanya. Selain itu kemungkinan adanya kebocoran soal pada saat dilakukan ujian, dinilai berbeda-beda oleh para mahasiswa, hal ini dipengaruhi dengan bagaimana pola mengajar para dosen; di mana ketika menjelang diadakannya ujian, dosen ataupun pengajar telah memberikan kisi-kisi atau detail materi yang sekiranya akan diujikan. Hal ini dirasa akan membantu mahasiswa dalam mendapatkan hasil yang baik namun di sisi lain juga membatasi lingkup penguasaan materi oleh mahasiswa secara maksimal. Kebocoran soal terkadang terjadi ketika melalui email internet, di mana salah satu mahasiswa diharapkan mendownload baru diperbanyak untuk dibagikan kepada seluruh mahasiswa. Artinya bahwa mahasiswa yang mengunduh tadi telah lebih dulu siap dengan berbagai jawabannya ketika jawaban tersebut merupakan jawaban non-alternatif (hanya beberapa yang benar) bukan pendapat pribadi. Tampak juga dalam data bahwa soal-soal yang dipakai dalam ujian lisan hampir sama kepada semua mahasiswa tetapi diujikan kepada mahasiswa yang berbeda pada waktu yang berbeda pula.

Data di lapangan menunjukkan bahwa nilai ujian yang diterima oleh mahasiswa cukup sesuai dengan kemampuan yang mereka miliki

dan berdasarkan pada penguasaan mereka akan materi. Hal ini sesuai dengan kondisi ketika mahasiswa memiliki penguasaan yang baik akan materi dan mampu menjawab soal yang diujikan, mahasiswa akan mendapatkan hasil serupa, namun juga berbanding sebaliknya. Di sisi lain, penilaian yang diberikan oleh dosen atau pengajar kepada mahasiswa tidak mutlak berdasarkan pada soal dan jawaban pada ujian, namun juga penilaian yang mampu mendukung selama proses belajar-mengajar berlangsung. Keaktifan dan antusiasme mahasiswa turut mempengaruhi hal tersebut. Dalam praktek pelaksanaan ujian mandiri, ketika ujian berupa tugas di luar waktu ujian terjadwal juga rentan akan keaslian pekerjaan mahasiswa. Hal ini semakin tidak reliable ketika mahasiswa dengan mudahnya mengandalkan pekerjaan tersebut kepada temannya dan ia mengumpulkannya sebagaimana hasil kemampuan dan usahanya sendiri; maka standar nilai yang diberikan dapat disimpulkan tidak sesuai karena tidak berdasarkan kompetensi yang dimiliki mahasiswa

### **Kepraktisan Soal Ujian**

Dilihat dari sisi kepraktisan tentunya ujian yang bergiliran berpotensi terjadi kebocoran soal meski sebenarnya bisa diatasi dengan soal yang berbeda, namun justru dosen akan kesulitan sebab harus membuat soal yang berbeda sejumlah mahasiswa. Banyak dosen yang memberikan ujian non-tes mengampu tiga kelas bersama. Dilihat dari sudut kepraktisan tentu ini tidak praktis jika ujian yang ada membutuhkan waktu per mahasiswa lima sampai lima belas menit menjadi tidak praktis.

Diungkapkan oleh mahasiswa waktu yang dibutuhkan untuk ujian non tulis yang amat beragam, mulai dari 2 menit bahkan hingga 3 bulan. Namun sekitar 11 mahasiswa mengikuti ujian non-tulis hanya 1 jam. Ini merupakan ujian non-tulis yang mengikuti jadwal akademik. Jika dilihat dari segi kepraktisan untuk 2 menit per-mahasiswa tentu akan sangat tidak praktis, terlebih jika jumlah mahasiswa banyak dengan asumsi dosen mengajar 3 kelas. Berdasarkan jumlah kelas, dilihat dari sudut kepraktisan tentu ujian non-tulis tidak praktis jika pelaksanaan ujian yang ada membutuhkan waktu per-mahasiswa 5 menit.

Beberapa mahasiswa berharap bahwa ujian hendaknya menggunakan ujian tertulis saja tidak perlu ujian non-tulis dan sebagian mahasiswa yang lain juga menghendaki ujian praktik dilaksanakan pada tempat yang sesuai. Jika masih menggunakan ujian non tulis diharapkan ada waktu khusus, ada waktu yang lebih dibandingkan ujian tertulis. Sebab ujian non tulis yang selama ini berjalan waktunya mepet. Selain itu mahasiswa juga berharap agar soal yang dibuat tidak terlalu sulit sehingga mahasiswa bisa mengerjakan soal dengan baik.

## **PENUTUP**

Hasil pembahasan tulisan ini mendapatkan simpulan sebagai berikut:

### a) Tingkat Validitas Soal yang Diujikan

Tingkat validitas soal dilihat dari butir-butir pertanyaannya cukup mengukur kompetensi mata kuliah yang direncanakan, namun soal ujian dalam pelaksanaan ujian juga berkembang sesuai situasi dan kondisi Dosen pada waktu itu sehingga mahasiswa seringkali mendapati soal ujiannya lebih sulit dari materi yang diajarkan. Hal ini disebabkan beberapa hal, antara lain mahasiswa kurang sungguh-sungguh dalam belajar. Soal ujian lisan yang diberikan menghendaki penalaran mahasiswa sendiri dikaitkan dengan kondisi keterkinian jauh dari materi yang telah disampaikan dosen saat proses pembelajaran.

Realitas ini mengharuskan dosen mengkaitkan materi perkuliahan dengan kondisi terkini sehingga ketika dibuat dalam soal ujian mahasiswa tidak mengalami kesulitan. Sebelum ujian perlu disampaikan secara umum kisi-kisi soal yang diujikan sesuai dengan materi perkuliahan yang diajarkan, dan dalam pelaksanaan ujian sebaiknya dihindari model soal yang tingkat kesulitan atau kompleksitas pemahaman yang dituntut dalam soal tes jauh lebih tinggi dari kompetensi yang telah disampaikan dalam kelas.

Tes lisan dengan cara tidak terjadwal serempak dan mandiri sangat baik ketika diujikan untuk menanyakan perbedaan-perbedaan pendapat atau judgment perorangan terhadap kasus tertentu. Tetapi ketika ujian tersebut menanyakan kompetensi referensial dan normatif sebagaimana dalam buku sumber, tes lisan terbukti kurang efektif untuk dilaksanakan. Dalam praktek pelaksanaan ujian lisan mandiri banyak mahasiswa yang

menjawab bahwa mereka mendapatkan soal yang sama untuk kelas paralel dan diujikan lagi pada kelas yang sama dengan waktu yang berbeda untuk mahasiswa lainnya. Artinya mahasiswa dapat mempersiapkan jawaban sebelum mengikuti ujian, karena mendapat bocoran soal dari kawannya. Hal tersebut bisa saja terjadi dan masuk akal karena seorang dosen tentunya sangat berat jika harus membuat ratusan soal yang berbeda untuk pelaksanaan ujian mandiri bagi ratusan mahasiswanya.

Tes lisan kemungkinan sangat tinggi bersifat terbuka terhadap isu-isu atau *hello effect* dari luar materi ujian yang tidak mengukur kompetensi mahasiswa. Sehingga konsistensi tujuan tes bisa berubah di tengah jalan karena faktor lingkungan tempat tes diselenggarakan.

#### b) Tingkat reliabilitas Soal yang Diujikan

Mahasiswa saling memberitahu soal-soal yang telah ditekankan kepada yang lebih awal, sehingga bobot kesulitan soalnya tidak merata bagi semua kelas yang diuji. Tingkat subyektifitas dosen sangat tinggi, hal ini ditunjukkan dengan ada perlakuan yang tidak samapada setiap mahasiswa yang mengikuti matakuliah tersebut. Pada kondisi yang sebaliknya, mahasiswa mengungkapkan bahwa perlakuan dosen saat menguji tidak adil karena dosen masih bersifat subjektif terhadap mahasiswa dan cenderung membedakan. Arah pertanyaan atau soal yang diujikan pun akan berubah-ubah karena sikap dosen yang subyektif, sehingga hasil yang diterima oleh mahasiswa akan berbeda dan tidak sesuai ketika dibandingkan dengan kompetensi yang dimilikinya.

#### c) Tingkat Kepraktisan Soal Ujian

Waktu yang dibutuhkan untuk ujian non tulis yang sangat beragam, mulai dari 2 menit hingga 3 bulan. Namun sebagian mahasiswa mengikuti ujian non tulis hanya 1 jam karena mengikuti jadwal akademik. Jika dilihat dari segi kepraktisan untuk 2 menit per-mahasiswa tentu akan sangat tidak praktis, terlebih jika jumlah dosen banyak dengan asumsi dosen mengajar 3 kelas. Keragaman ujian non tulis, ada yang dilaksanakan secara bergiliran ada pula yang bersama-sama dengan jumlah peserta mencapai 30 mahasiswa. Dilihat dari sisi kepraktisan, ujian bergiliran berpotensi terjadi kebocoran soal meski dapat diatasi dengan soal yang berbeda, namun dosen akan kesulitan ketika harus membuat soal sebanyak jumlah mahasiswa. Ujian non-tulis

yang diikuti mahasiswasesuai jadwal akademik hanya berkisar sekitar 60–90 menit saja. Ujian non-tes yang mengikuti jadwal akademik ini berpotensi tidak praktis. Begitu pula ujian non-tulis yang dilaksanakan 2 hari berturut turut selama 4 Minggu. Ujian semacam ini jelas tidak praktis, hampir 4 pertemuan yang seharusnya menjadi jam kuliah justru hanya digunakan untuk menguji mahasiswa. Soal ujian yang diberikan belum menunjukkan tingkat kepraktisan yang baik, terbukti dengan beberapa pendapat mahasiswa yang mengatakan bahwa ujian non-tulis tersebut kurang efektif.

## DAFTAR PUSTAKA

- Sudjiono, Anas. *Pengantar Evaluasi Pendidikan*. Jakarta : RajaGrafindo Persada, 2006.
- Arifin, Zainal. *Evaluasi Pembelajaran Prinsip, Teknik, Prosedur*. Bandung: Remaja Rosdakarya, 2010.
- Arikunto, Suharsimi dan Safruddin AbdulJabar. *Evaluasi Program Pendidikan Pedoman Praktis Bagi Mahasiswa dan Praktisi Pendidikan*. Jakarta: Bumi Aksara, 2010.
- Dimiyati dan Mudjiono. *Belajar dan Pembelajaran*. Jakarta: Rineke Cipta, 2006.
- Purwanto, Ngalim. *Prinsip-Prinsip dan Teknik Evaluasi Pengajaran*. Bandung: Remaja Rosdakarya, 2010.
- A.Zainul. *Alternatif Assesment*. Jakarta: Dirjen Dikti, 2001.
- “Evaluasi Pendidikan”. [http://www.scribd.com/doc/30098749/Pengertian-Evaluasi-Pendidikandiakses tanggal 3 Agustus 2012](http://www.scribd.com/doc/30098749/Pengertian-Evaluasi-Pendidikandiakses%20tanggal%203%20Agustus%202012).
- Kasmo, Sukasmo “Makna Evaluasi Pendidikan”. dalam <http://edukasi.kompasiana.com/2012/03/05/makna-evaluasi-pendidikan> diakses tanggal 3 Agustus 2012.