

DEVELOPMENT STRATEGY OF INSTRUMENT TEST ASSESSMENT FOR THE HIGHER ORDER THINKING STUDENTS

Laili Etika Rahmawati
Siti Kholifatul Hasanah
Anggraeni Dwi Sulistyowati
Indonesian Language Education Faculty of Teacher and Training Education
Universitas Muhammadiyah Surakarta
Laili.Rahmawati@ums.ac.id

ABSTRACT: The mid- and final-semester examinations are part of the lecturing evaluation. The evaluation technique used for assessing the students' competency is usually based on a test. As an assessment, the test can be developed in the subjective and objective types, but each type has different characteristics. It can be useful for assessing the critical thinking students' who not only memorize and understand the materials of the test. The research paper aims to describe a development strategy of instrument test assessment of the critical thinking students. It is the introductory study result of the mid-semester examination for Learning Evaluation course. Based on the identification of questions, answer sheet, and students' responses when evaluating the peer-students, it is stated that the instrument test development can be helpful for assessing the critical thinking students' answers. It is proved that there were the heterogonous answers based on the scheme that was a cheat-on-exam result for their friends and books.

Keywords: *instrument test, assessment, competency, higher order thinking*

ABSTRAK: Ujian tengah semester dan ujian akhir semester merupakan bagian dari evaluasi perkuliahan. Teknik penilaian yang digunakan untuk mengukur kompetensi mahasiswa biasanya berbentuk tes. Tes yang disajikan sebagai alat ukur kompetensi mahasiswa dapat dikembangkan dalam bentuk tes subjektif maupun tes objektif. Masing-masing jenis tes mempunyai kelebihan dan kelemahan. Tes yang dikembangkan sebagai alat ukur kompetensi mahasiswa hendaknya dapat digunakan untuk mengukur kompetensi berpikir tingkat tinggi, bukan sekadar menghafal dan memahami materi yang diujikan. Makalah ini bertujuan untuk mendeskripsikan strategi pengembangan instrumen tes pengukur kompetensi berpikir tingkat tinggi mahasiswa. Makalah ini merupakan hasil dari studi pendahuluan terhadap instrumen tes UTS mata kuliah Evaluasi Pembelajaran. Berdasarkan hasil identifikasi yang dilakukan terhadap soal, lembar jawab, serta respon mahasiswa saat melakukan penilaian sejawat diketahui bahwa instrumen tes yang dikembangkan untuk mengukur kompetensi mahasiswa penempuh mata kuliah Evaluasi Pembelajaran secara efektif dapat mengukur kompetensi berpikir tingkat tinggi mahasiswa. Hal ini ditunjukkan dengan adanya keragaman jawaban mahasiswa yang didasarkan pada skemata yang dimiliki bukan dari hasil menyontek dari teman ataupun buku catatan.

Kata Kunci: instrumen tes, pengukur, kompetensi, berpikir tingkat tinggi

PENDAHULUAN

Ujian tengah semester dan ujian akhir semester merupakan bagian dari evaluasi perkuliahan. Teknik penilaian yang digunakan untuk mengukur kompetensi mahasiswa biasanya berbentuk tes. Tes yang disajikan sebagai alat ukur kompetensi mahasiswa dapat dikembangkan dalam bentuk tes subjektif maupun tes objektif. Masing-masing jenis tes mempunyai kelebihan dan kelemahan. Tes yang dikembangkan sebagai alat ukur kompetensi mahasiswa hendaknya dapat digunakan untuk mengukur kompetensi berpikir

tingkat tinggi, bukan sekadar menghafal dan memahami materi yang diujikan.

Bentuk tes yang dipilih selalu disesuaikan dengan tingkat kemampuan peserta didik yang akan diukur. Bentuk tes memilih jawaban benar salah, isian singkat, dan menjodohkan merupakan alat yang hanya menilai kemampuan berpikir rendah, yaitu kemampuan mengingat. Pilihan ganda dan sebab akibat dapat digunakan untuk menilai kemampuan siswa dalam mengingat dan memahami. Adapun tes uraian dapat digunakan untuk menilai kemampuan mengingat, memahami,

mengorganisasikan gagasannya atau hal-hal yang sudah dipelajari. Tes uraian disebut juga dengan tes esai adalah suatu bentuk pertanyaan yang menuntut jawaban dalam bentuk uraian dengan mempergunakan bahasanya sendiri (Nurgiyantoro, 2001:71).

Masing-masing bentuk tes memiliki keunggulan dan kelemahan. Nurgiyantoro (2001:72) menyebutkan keunggulan tes bentuk uraian adalah: (1) tes esai tepat untuk menilai proses berpikir yang melibatkan kognitif tinggi; (2) tes esai memberi kesempatan siswa untuk mengemukakan jawabannya ke dalam bahasa yang runtut sesuai dengan gayanya; (3) tes esai memaksa siswa untuk berpikir sendiri, tidak bisa untung-untungan; dan (4) tes esai mudah disusun sehingga tidak menghabiskan waktu. Selain keunggulan tersebut, tes esai memiliki kelemahan yaitu: (1) kadar validitas dan reliabilitasnya rendah; (2) akibat terbatasnya bahan yang diteskan, dapat terjadi hasil yang bersifat kebetulan; (3) standar penilaiannya sukar ditentukan; dan (4) waktu untuk memeriksa jawaban siswa relatif lama.

Tes memilih jawaban (objektif) juga memiliki keunggulan dan kelemahan. Tes Objektif memiliki keunggulan antara lain: (1) bahan yang diteskan dapat meliputi ruang lingkup yang luas, (2) reliabilitasnya tinggi karena jawaban yang benar hanya satu, (3) mudah dikoreksi karena bisa diwakilkan atau bahkan diwakilkan komputer, dan (4) koreksi dapat dikerjakan dengan cepat, bahkan sangat cepat (Nurgiyantoro, 2001:76-77). Di samping keunggulan tersebut, tes objektif memiliki kelemahan: (1) penyusunan tes objektif membutuhkan waktu yang lama, (2) sering bahan tes tidak komprehensif, (3) tingkatan kognitif yang diungkap hanya pada tingkatan dasar, (4) siswa sering menjawab dengan asal-asalan, atau untung-untungan, (5) membutuhkan waktu dan biaya yang banyak dalam penggandaannya (Nurgiyantoro, 2001:78-79).

Dalam menentukan bentuk tes dan menyusun instrumennya perlu mempertimbangkan berbagai hal. Hal-hal itu antara lain: (1) karakteristik mata pelajaran, (2) keluasan ruang lingkup materi yang akan diuji, (3) materi berkaitan dengan kesesuaian soal dengan standar kompetensi, kompetensi dasar, dan indikatornya, (4) konstruksi soal harus jelas, dan (5) bahasa yang digunakan tidak menimbulkan penafsiran ganda (Puskur Depdiknas, 2006: 15).

Bentuk tes uraian merupakan bentuk tes yang dianjurkan dalam penilaian autentik dibanding dengan bentuk penilaian objektif. Tes uraian ini

dapat menilai berbagai jenis kompetensi, misalnya: mengemukakan pendapat, berpikir logis, dan menyimpulkan. Dengan hasil yang seperti itu, guru dapat menentukan tindak lanjut dengan efektif.

Berdasarkan taksonomi Bloom yang telah direvisi proses kognitif dibagi menjadi kemampuan berpikir tingkat rendah dan kemampuan berpikir tingkat tinggi. Kemampuan yang termasuk dalam kategori kemampuan berpikir tingkat rendah yaitu mengingat, memahami, dan menerapkan, sedangkan kemampuan yang masuk kategori kemampuan berpikir tingkat tinggi yaitu kemampuan menganalisis, mengevaluasi, dan menciptakan (Anderson dan Krathwohl, 2001:30).

Setiawan, dkk. (2014:244) menyatakan bahwa keterampilan berpikir tingkat tinggi merupakan kemampuan menghubungkan, memanipulasi, dan mentransformasi pengetahuan serta pengalaman yang sudah dimiliki untuk berpikir secara kritis dan kreatif dalam upaya menentukan keputusan dan memecahkan masalah pada situasi baru. Seorang siswa dikatakan mampu menyelesaikan masalah apabila ia dapat menerapkan pengetahuan yang telah diperoleh sebelumnya ke dalam situasi baru yang belum dikenal. Kemampuan inilah yang biasa dikenal sebagai keterampilan berpikir tingkat tinggi. Hellsten (2010:654) menyatakan bahwa ada lima hal yang dapat dijadikan acuan untuk mengukur kemampuan berpikir kritis, yaitu: (1) pengetahuan awal (skemata); (2) kriteria untuk menentukan sesuatu; (3) kosakata berpikir kritis; (4) strategi berpikir; dan (5) kebiasaan berpikir.

Nitko dan Brookhart (2011:223) menjelaskan bahwa ketentuan dasar penilaian kemampuan berpikir tingkat tinggi adalah menggunakan tugas-tugas yang memerlukan penggunaan pengetahuan dan keterampilan dalam situasi baru. Untuk melakukan penilaian terhadap kemampuan berpikir tingkat tinggi harus menggunakan bahan-bahan baru. Salah satu cara yang dapat dilakukan adalah menggunakan set-set item yang bergantung pada konteks.

Berdasarkan latar belakang masalah di atas perlu dilakukan kajian tentang strategi pengembangan instrumen tes pengukur kompetensi berpikir tingkat tinggi mahasiswa. Rumusan ini relevan dengan penelitian-penelitian yang telah dilakukan oleh peneliti terdahulu khususnya dalam pengembangan instrumen tes bidang eksakta.

Peneclitian yang dilakukan oleh Istiyono, Mardapi, dan Suparno (2014) yang berjudul *Pengembangan Tes Kemampuan Berpikir Tingkat*

Tinggi Fisika (PsysTHOTs) Peserta Didik SMA menemukan bahwa (1) instrumen *PhysTHOTS* dikembangkan dalam bentuk pilihan ganda beralasan pada kemampuan menganalisis, mengevaluasi, dan menciptakan untuk materi fisika gerak, gaya, usaha dan energi, serta momentum dan impuls yang terdiri atas perangkat tes A dan perangkat tes B yang masing-masing terdapat 26 *item* dengan 8 *anchor item*; (2) instrumen *PhysTHOTS* telah memenuhi validitas isi dengan *expert judgment* dan telah mendapatkan bukti empiris validitas konstruk *fit* pada *Partial Credit Model (PCM)* berdasarkan data politomus empatkategoris; (3) seluruh *item* pada *PhysTHOTS* dalam kriteria baik karena tingkat kesulitannya berada pada rentang antara -2,00 sampai dengan 2,00. Reliabilitas *PhysTHOTS* telah memenuhi syarat, bahkan termasuk tinggi (koefisien reliabilitas lebih dari 0,90). Berdasarkan fungsi informasi, *PhysTHOTS* sangat tepat digunakan untuk mengukur kemampuan berpikir tingkat tinggi fisika peserta didik yang berkemampuan dari -0,80 sampai 3,40.

Penelitian tersebut mengindikasikan bahwa perlu strategi pengembangan instrumen secara khusus untuk dapat mengukur kemampuan berpikir tingkat tinggi. Kelemahan tes objektif (dalam konteks ini pilihan ganda) yang selama ini dikategorikan ke dalam jenis tes pengukur kompetensi berpikir tingkat rendah dapat diminimalisasi dengan strategi pengembangan model *PhysTHOTS* dengan langkah-langkah pengembangan yang tentunya tidak sederhana.

Dalam bidang bahasa, penelitian yang dilakukan Soleimani, H. dan Kheiri, S. (2016) dengan judul *An Evaluation of TEFL Postgraduates' Testing Classroom Activities and Assignments Based on Bloom's Revised Taxonomy* menyimpulkan bahwa Taksonomi Bloom versi revisi digunakan sebagai acuan pengukur kompetensi berpikir tingkat rendah, menengah, dan tinggi. Taksonomi Bloom dapat digunakan sebagai penyeimbang antara penilaian dan pertanyaan evaluasi dalam kelas. Berdasarkan hasil penelitian yang dilakukan direkomendasikan untuk meningkatkan aktivitas pembelajaran dan penilaian dengan mengacu pada Taksonomi Bloom dalam mengembangkan kurikulum, mengembangkan instrumen tes, dan menilai kompetensi mahasiswa.

Penelitian Eric Zhi-Feng, Yi-Chin Zhuo dan Shyan-Ming Yuan (2004) yang berjudul *Assesing Higher -Order Thinking Using A Networked Portfolio System With Peer Assesment* menemukan bahwa salah satu upaya yang dapat

dilakukan untuk meningkatkan keterampilan berpikir tingkat tinggi siswa dan mahasiswa adalah menerapkan teknik penilaian portofolio dan penilaian teman sejawat. Penerapan teknik penilaian portofolio dan penilaian teman sejawat dapat merekam aktivitas-aktivitas selama proses pembelajaran. Siswa cenderung berperan aktif dan mempunyai kesempatan untuk mengetahui kompetensinya dengan hasil review portofolio oleh teman sejawat. Komentar yang disampaikan oleh teman sejawat dapat dijadikan tolok ukur kompetensi berpikir tingkat tinggi mahasiswa.

METODE PENELITIAN

Metode yang digunakan dalam penelitian ini adalah deskriptif kualitatif. Objek penelitian ini adalah instrumen tes ujian tengah semester mata kuliah Evaluasi Pembelajaran yang ditempuh oleh mahasiswa semester IV Program Studi Pendidikan Bahasa FKIP UMS semester genap 2015/2016. Teknik pengumpulan data yang digunakan dalam penelitian ini adalah dokumentasi dan wawancara. Dokumen yang dianalisis dalam penelitian ini adalah instrumen tes UTS mata kuliah Evaluasi pembelajaran dan respon mahasiswa terhadap instrumen tes tersebut. Instrumen tes dianalisis berdasarkan teori-teori tentang kompetensi berpikir tingkat tinggi.

HASIL DAN PEMBAHASAN

Mata kuliah Evaluasi Pembelajaran merupakan mata kuliah hasil peninjauan kurikulum yang dilakukan oleh Program Studi PBI mengacu pada KKNi dan SNPT yang semula memiliki beban 4 sks disederhanakan menjadi 2 sks dengan target capaian pembelajaran yang relatif sama. Materi perkuliahan yang harus dikuasai oleh mahasiswa sampai UTS adalah tentang hakikat evaluasi, penilaian autentik, teknik-teknik penilaian autentik, dan pengembangan instrumen penilaian. Agar mampu mengukur kompetensi mahasiswa berdasarkan capaian pembelajaran yang telah ditetapkan maka perlu strategi khusus untuk mengembangkan instrumen tes untuk mengukur kompetensi mahasiswa.

Mata kuliah dengan bobot 2 sks diberikan waktu mengerjakan maksimal 90 menit. Berdasarkan pertimbangan tersebut maka soal UTS dikembangkan menjadi 4 butir soal dengan bobot nilai yang berbeda-beda disesuaikan dengan tingkat kesulitan dan tingkatan kompetensi berpikir.

Berikut soal UTS mata kuliah Evaluasi Pembelajaran yang telah dikembangkan. 243

Soal nomor 1

Deskripsikan hubungan antara evaluasi, penilaian, pengukuran, dan tes! Berdasarkan deskripsi tersebut tentukan pernyataan berikut benar atau salah!

- a. *Penilaian merupakan bagian dari evaluasi.*
 - b. *Pengukuran dan penilaian mempunyai makna yang sama dalam konteks evaluasi pembelajaran bahasa.*
 - c. *Pengukuran merupakan kegiatan untuk menentukan nilai.*
 - d. *Istilah measurement sering juga disebut dengan penilaian.*
 - e. *Penilaian harus dilaksanakan dengan teknik tes.*
- (skor maksimal 20)*

Soal nomor 1 bertujuan untuk mengukur kompetensi mahasiswa berkaitan materi hakikat evaluasi. Capaian pembelajaran yang harus dikuasai oleh mahasiswa adalah mahasiswa mampu memahami hakikat evaluasi dengan memahami perbedaan konsep evaluasi, penilaian, pengukuran, dan tes. Strategi penyajian soal menggunakan jenis tes objektif benar-salah. Secara teoretis salah satu kelemahan tes objektif adalah hanya mampu mengukur kemampuan kognitif tingkat rendah. Namun, dengan strategi mendeskripsikan terlebih dahulu maka mahasiswa mau tidak mau harus memahami konsep dasar evaluasi, penilaian, pengukuran, dan tes. Dengan demikian gambaran hubungan konsep tersebut dapat dijadikan dasar untuk menjawab soal objektif benar salah poin a, b, c, d, dan e.

Soal nomor 2

Bilamana penilaian dikatakan autentik? (skor maksimal 10)

Soal nomor 2 dimanfaatkan oleh penguji untuk menilai kompetensi berpikir tingkat rendah. Hal ini dilakukan untuk memberikan peluang mahasiswa untuk bisa menjawab benar. Dasar penyajian soal tersebut adalah dalam instrumen tes yang disajikan disarankan tingkat kesulitan soal beragam. Sekalipun soal tersebut masuk kategori tes pengukur kompetensi berpikir tingkat rendah karena jawaban masih berorientasi pada kemampuan siswa menghafal materi perkuliahan, tetapi penyajian jawaban mahasiswa bisa dijadikan indikator kompetensi berpikir tingkat tinggi. Hal ini didukung oleh teori Hellsten (2010:654) yang menyatakan bahwa kosakata yang digunakan merupakan salah satu indikator kemampuan berpikir seseorang.

Soal nomor 3

Bu Xenia adalah guru Mata Pelajaran Bahasa Indonesia di SMA Daihatsu. Pada pertemuan ke-3 Bu Xenia melaksanakan penilaian keterampilan berbicara siswa. Instrumen yang digunakan untuk menilai keterampilan berbicara siswa berupa daftar cek. Bu Xenia menentukan kriteria penilaian baik mendapat skor 1 dan tidak baik mendapat skor 0. Berdasarkan kriteria tersebut apa perbedaan antara siswa yang belum maju dengan siswa yang sudah maju tetapi semua aspek yang dinilai berkategori tidak baik. Saran apa yang Saudara usulkan agar perbedaan antara siswa yang belum maju dengan yang sudah maju dengan kategori tidak baik benar-benar dapat dibedakan? (skor maksimal 30)

Soal nomor 3 bertujuan untuk mengukur kompetensi mahasiswa khususnya berkaitan dengan penggunaan instrumen penilaian berupa daftar cek, *rating scale*, dan penilaian fleksibel. Kasus penilaian diambil dari permasalahan yang sangat umum di masyarakat. Secara umum masyarakat memahami bahwa nol (0) dan kosong (Ø) bermakna sama. Sebagai contoh ketika seseorang mengisi pulsa dan ditanya nomor telepon genggamnya sering menjawab kosong delapan lima enam empat tujuh kosong dua kosong tiga enam tujuh (085647020367). Kata kosong dalam nomor telepon genggam tersebut seharusnya dituturkan dengan kata nol. Ternyata permasalahan ini juga berdampak pada pelaksanaan penilaian di sekolah. Banyak guru yang beranggapan bahwa nol dan kosong merupakan kata yang bersinonim dan bisa digunakan secara bergantian.

Berdasarkan permasalahan tersebut maka soal nomor 3 disajikan untuk mengukur kompetensi berpikir tingkat tinggi mahasiswa dengan dua strategi penyajian, yaitu konteks dan pengetahuan awal (skemata). Strategi penyajian yang mengacu pada konteks ditunjukkan dengan nama guru dan nama sekolah yang tidak lazim tetapi dikenal baik oleh mahasiswa. Berdasarkan hasil wawancara dengan mahasiswa strategi ini memberikan dampak mahasiswa ketika membaca soal sangat tertarik dan serius untuk memahami pertanyaan yang disajikan. Strategi yang kedua adalah skemata yaitu untuk mengembalikan konsep yang salah agar menjadi benar.

Soal nomor 3 dijawab beragam oleh mahasiswa. Keberagaman jawaban tersebut menunjukkan bahwa mahasiswa mempunyai persepsi yang berbeda juga terhadap pertanyaan

yang disajikan. Adapun jawaban yang muncul dari pertanyaan tersebut antara lain sebagai berikut.

- 1) Perbedaan antara siswa yang belum maju dengan yang sudah maju adalah siswa yang belum maju nilainya kosong siswa yang sudah maju nilainya nol. Saran yang diajukan adalah guru mengubah kriteria penilaian dari baik 1 dan tidak baik 0 menjadi baik 2 dan tidak baik 1.
- 2) Perbedaan antara siswa yang belum maju dengan yang sudah maju adalah siswa yang belum maju belum mendapat nilai dan siswa yang sudah maju sudah mendapat nilai. Saran yang diberikan adalah siswa yang belum maju daftar nilainya distabilo.
- 3) Siswa yang sudah maju dengan siswa yang belum maju nilainya sama-sama nol. Saran: guru memberikan nilai keaktifan siswa yang sudah maju tetapi nilainya tidak baik (jangan diberi nilai 0 karena sudah maju) dan meminta siswa yang belum maju untuk maju minggu depan.
- 4) Siswa yang belum maju belum mendapat nilai (kosong), sedangkan siswa yang sudah maju tetapi tidak baik mendapat nilai nol (0). Agar mudah dibedakan antara siswa yang sudah maju tetapi nilainya tidak baik dengan siswa yang belum maju maka ada dua alternatif yang bisa dilakukan, yaitu: (1) tetap menggunakan daftar cek dengan mengubah kriteria penilaian dari baik =1 menjadi 2 dan yang tidak baik = 0 menjadi 1; (2) mengubah instrumen penilaian menjadi *rating scale* atau skala penilaian fleksibel.

Jawaban tersebut mengindikasikan bahwa soal yang disajikan mampu menggali kompetensi berpikir tingkat tinggi mahasiswa dengan adanya keragaman jawaban yang disajikan. Hal ini sesuai dengan teori Hellsten (2010:654) menyatakan bahwa ada lima hal yang dapat dijadikan acuan untuk mengukur kemampuan berpikir kritis, yaitu: (1) pengetahuan awal (skemata); (2) kriteria untuk menentukan sesuatu; (3) kosakata berpikir kritis; (4) strategi berpikir; dan (5) kebiasaan berpikir. Mendukung pernyataan tersebut, Shouhong Wang & Hai Wang. (2012) dalam penelitiannya tentang kontribusi skemata terhadap kompetensi berpikir tingkat tinggi menemukan bahwa skemata merupakan unsur penting dalam penyajian pendapat yang mampu mendeskripsikan kompetensi berpikir tingkat tinggi.

Soal nomor 4

Pak Pajero adalah guru Mata Pelajaran Bahasa Indonesia di SMP Mitsubishi. Pak Pajero pada pertemuan ke-5 mengajarkan

kompetensi dasar menyajikan teks tanggapan deskriptif dalam bahasa Indonesia lisan dan tulis untuk menyusun kembali dan membuat tanggapan pribadi atas karya budaya masyarakat Indonesia yang penuh makna secara mandiri. Salah satu indikator yang harus dicapai siswa adalah menyajikan teks tanggapan deskriptif lisan (12 s.d. 15 kalimat) tentang karya budaya masyarakat Indonesia yang penuh makna menurut tahapan, kohesi dan leksikogramatika yang baik dan benar dan dengan bantuan media (gambar, diagram, dan video).

Berdasarkan ilustrasi tersebut jawablah pertanyaan berikut ini.

- a. ***Teknik penilaian autentik apa yang tepat digunakan oleh Pak Pajero untuk mengetahui pencapaian hasil belajar siswa terhadap indikator pembelajaran yang telah ditentukan?***
- b. ***Buatlah instrumen penilaian dalam bentuk rating scale dan skala penilaian fleksibel beserta simulasi sistem penilaiannya!***

Soal nomor 4 dibuat lebih kompleks dibandingkan dengan soal nomor 1, 2, dan 3. Dalam soal nomor 4 mahasiswa dituntut untuk mampu menentukan teknik penilaian yang relevan berdasarkan konteks cerita, menentukan aspek-aspek yang dinilai, dan memproduksi instrumen penilaiannya. Kompetensi menentukan teknik penilaian yang tepat secara tidak langsung mampu menguji penguasaan mahasiswa khususnya tentang teknik-teknik penilaian autentik. Aspek-aspek yang dinilai harus ditentukan berdasarkan deskripsi dalam soal ujian. Namun, faktanya mahasiswa banyak yang menjawab berdasarkan asumsi pribadi dengan mengabaikan konteks dalam teks soal. Aspek-aspek yang dinilai sebagai dasar pengembangan instrumen dinyatakan secara eksplisit di dalam soal, yaitu: (1) panjang kalimat (12 s.d. 15 kalimat), (2) tahapan penyajian, (3) kohesi dan leksikogramatika, dan (4) media. Namun, mahasiswa yang belum menguasai strategi penentuan aspek penilaian banyak yang terjebak menentukan aspek penilaian kompetensi berbicara yang diajarkan saat perkuliahan. Risiko dari kesalahan penentuan aspek penilaian berdampak pada kesalahan pengembangan instrumen berupa *rating scale* dan skala penilaian fleksibel yang merupakan jawaban dari pertanyaan soal nomor 4b.

Hal ini sesuai dengan teori yang disampaikan oleh Nitko dan Brookhart (2011:223) bahwa ketentuan dasar penilaian kemampuan ber

tingkat tinggi adalah menggunakan tugas-tugas yang memerlukan penggunaan pengetahuan dan keterampilan dalam situasi baru, menggunakan bahan-bahan baru, dan set-set item yang bergantung pada konteks.

SIMPULAN

Berdasarkan pada hasil dan pembahasan di atas dapat disimpulkan bahwa pengembangan instrumen tes pengukur kompetensi berpikir tingkat tinggi mahasiswa dapat dilakukan dengan beberapa strategi yaitu mengacu pada: (1) pengetahuan awal (skemata); (2) kriteria untuk menentukan sesuatu; (3) kosakata berpikir kritis; (4) strategi berpikir; dan (5) kebiasaan berpikir. Selain itu dasar penilaian kemampuan berpikir tingkat tinggi adalah menggunakan tugas-tugas yang memerlukan penggunaan pengetahuan dan keterampilan dalam situasi baru, menggunakan bahan-bahan baru, dan set-set item yang bergantung pada konteks. Dengan strategi tersebut, kecurangan dalam pelaksanaan tes dapat diminimalisasi karena cara penyajian jawaban dapat dijadikan indikasi apakah mahasiswa menyontek atau mengerjakan sendiri.

DAFTAR PUSTAKA

- Anderson, L.W., and Krathwohl, D.R. (2001). *A Taxonomy of Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives*. New York: Longman.
- Eric Zhi-Feng, Yi Chin Zhuo, dan Shyan-Ming Yuan. 2004. "Assessing Higher Order Thinking Using a Networked Portfolio System with Perr Assesment". *International Journal of Instructional Media* 31 (2) page 139-149.
- Hellsten, Laurie-ann. 2010. "Critical Thinking Education and Assessment: Can Higher Order Thinking Be Tested?". *Canadian Journal of Education* 33 (3) 652 – 656.
- Istiyono, Edi, Mardapi, Djemari, Suparno. (2014). "Pengembangan Tes Kemampuan Berpikir Tingkat Tinggi Fisika (PsysTHOTs) Peserta Didik SMA". *Jurnal Pendidikan dan Evaluasi Pendidikan* 18 (1) 1-12.
- Nitko, A.J & Brookhart, S. M. (2011). *Educational assessment of students*. Boston: Pearson Education, Inc.
- Nurdiyantoro, Burhan. (2001). *Penilaian dalam Pengajaran Bahasa dan Sastra*. Yogyakarta: BPF.

- _____. (2010). *Penilaian Pembelajaran Bahasa Berbasis Kompetensi*. Yogyakarta: BPF.
- Pusat Kurikulum Badan Penelitian dan Pengembangan Depdiknas. (2006). *Model Penilaian Kelas Kurikulum Tingkat Satuan Pendidikan*. Jakarta: Depdiknas.
- Setiawan, Harianto, Dafik, dan Lestari, N.D.S. 2014. "Soal Matematika Dalam PISA Kaitannya dengan Literasi Matematika dan Keterampilan Berpikir Tingkat Tinggi". *Prosiding Seminar Nasional Matematika, Universitas Jember, 19 November 2014 halaman 244-251*.
- Shouhong Wang & Hai Wang. (2012). "Organizational Schemata of E-Portfolios for Fostering Higher-Order Thinking". *Inf Syst Front* 14:395–407.
- Soleimani, H. Dan Kheiri, S. (2016). "An Evaluation of TEFL Postgraduates' Testing Classroom Activities and Assignments Based on Bloom's Revised Taxonomy". *Theory and Practice in Language Studies, Vol. 6, No. 4, pp. 861-869, April 2016*