

ANALISIS TES PENILAIAN PENCAPAIAN KOMPETENSI PADA MAHASISWA KEBIDANAN

Ni Ketut Erawati¹

¹Program Studi D3 Kebidanan, Universitas Pendidikan Ganesha
Singaraja, Indonesia

e-mail: erawatitrinira@yahoo.com

Abstrak

Salah satu bentuk instrumen untuk mengukur hasil belajar adalah tes. Tes dikatakan baik apabila memenuhi syarat diantaranya valid dan reliabel, namun analisis tes masih jarang dilakukan. Penelitian ini bertujuan menganalisis perangkat tes penilaian pencapaian kompetensi (PPK) bidan. Analisis tes meliputi analisis tingkat kesukaran, daya pembeda, validitas dan reliabilitas tes. Penelitian dilakukan pada mahasiswa kebidanan semester V. Penelitian ini berjenis deskriptif kuantitatif, analisis data menggunakan rumus tingkat kesukaran, daya beda, poin biserial dan KR-20. Sampel berupa perangkat tes penilaian kompetensi yang terdiri dari 120 butir soal pilihan ganda. Hasil penelitian menunjukkan 57% butir soal memiliki tingkat kesukaran terkategori mudah, 33% memiliki daya pembeda terkategori jelek, 64% butir soal tidak valid artinya masih banyak butir soal yang harus diperbaiki atau harus dibuang. Uji reliabilitas pada butir soal yang valid menunjukkan tingkat reliabilitas sangat baik yaitu 95%. Dengan melakukan analisis butir soal harapannya akan mampu mengukur kompetensi mahasiswa sebenarnya.

Kata-kata kunci: tes penilaian kompetensi, validitas, reliabilitas.

Abstract

An examination test is one method to assess learning outcomes. Item analysis may be used to assess validity and reliability of the test. However, limited analysis is employed in the current practice. This research aims to analyse midwifery students' competency test. The analysis of Competency assessment tools includes level of difficulties and level of differences. This study was conducted on midwifery students in their fifth semester. The tools consist of 120 items with multiple choice questions. Students' response was analysed based on level of differences, biserial point and KR-20. Data was reported descriptively. This study showed that 58% items were at easy level, 33% items had low level of differences, and 64% items were invalid. This meant that many items in the tools need to be amended. Reliability test on valid items shows a very good level of reliability, which is 95%. By carrying out item analysis the expectation will be able to measure actual student competencies

Keywords : *competency assesment test, validity, reliability.*

PENDAHULUAN

Evaluasi hasil belajar dilakukan oleh pendidik untuk memantau proses, kemajuan, pencapaian dan perbaikan hasil belajar peserta didik secara berkesinambungan (Sukardi, 2011). Evaluasi hasil belajar mahasiswa biasanya diujikan melalui ujian tengah semester (UTS), ujian akhir semester (UAS) pada akhir pembelajaran. Tidak terkecuali pada mahasiswa Jurusan kebidanan. Sebagai salah satu pendidikan tinggi vokasi yang mengedepankan penguasaan keterampilan disamping juga ranah pengetahuan dan sikap, maka ketika mahasiswa akan naik tingkat ke jenjang berikutnya mahasiswa wajib untuk dinilai kemampuan kompetensinya secara bertahap.

Salah satu model evaluasi yang diterapkan di Jurusan Kebidanan adalah evaluasi melalui pelaksanaan Penilaian Pencapaian Kompetensi (PPK). PPK ini sebenarnya sudah berlangsung dari tahun 2015, namun pelaksanaannya belum pernah dievaluasi. Evaluasi PPK ini diselenggarakan secara bertahap, yakni PPK I pada akhir semester 2, PPK II yang dilakukan pada akhir semester 5 dan PPK III yang dilakukan pada akhir semester 6. Nilai PPK merupakan gambaran penguasaan kompetensi yang dipelajari mahasiswa dalam menempuh proses pembelajaran selama empat semester, sehingga diperlukan soal yang berkualitas baik. Evaluasi PPK ini dilakukan secara teoretik maupun keterampilan oleh pendidik. Adapaun bentuk evaluasi secara teori dilakukan dalam bentuk tes kompetensi objektif atau pilihan ganda. Secara umum kita ketahui bahwa fungsi tes secara umum yaitu (1) sebagai alat pengukur terhadap perkembangan atas kemajuan peserta didik, dan (2) sebagai alat ukur

pengukur keberhasilan program pengajaran (Sudijono, 2013). Seperti layaknya syarat tes pada umumnya, perangkat soal yang digunakan pun harus memenuhi unsur validitas dan reliabilitas. Perangkat tes yang digunakan selama ini untuk evaluasi sebenarnya sudah diujikan secara kualitatif melalui review butir soal yang dilakukan oleh pakar dibidangnya, namun secara kuantitatif evaluasi perangkat tes teoretik PPK ini belum pernah dilakukan. Hal ini tentunya menjadi suatu pemikiran bahwa apapun hasil evaluasi dari tes PPK yang dilakukan terhadap peserta didik akan dipengaruhi oleh kualitas tes yang digunakan. Namun kenyataannya dari hasil evaluasi pengetahuan yang telah dilakukan didapatkan nilai kelulusan mahasiswa masih rendah, hal ini dapat dilihat dari 78 orang peserta uji tulis PPK I, 40 orang diantaranya dinyatakan tidak lulus, 37 orang dinyatakan lulus dan 1 orang tidak hadir. Oleh karenanya banyak mahasiswa yang harus mengikuti remedial I bahkan sampai remedial II. Seperti kita ketahui bahwa keberhasilan seorang mahasiswa untuk mencapai prestasinya tidak saja dilihat dari kemampuannya sendiri tetapi juga dari kualitas perangkat evaluasinya.

Tes sebagai salah satu instrumen evaluasi tentu mengambil peran penting dalam hal ini. Kualitas tes sebagai salah satu alat evaluasi yang penting untuk diperhatikan. Kualitas tes dapat menjadi pedoman untuk penilaian ketepatan hasil belajar peserta didik. Pendidik perlu melakukan analisis terlebih dahulu sebelum melakukan penilaian hasil belajar.

Analisis merupakan kemampuan seseorang untuk merinci atau menguraikan suatu bahan atau keadaan menurut bagian-bagian yang lebih kecil

dan mampu memahami hubungan di antara bagian-bagian atau faktor-faktor yang satu dengan faktor-faktor lainnya (Sudijono, 2013). Analisis yang perlu dilakukan pendidik adalah analisis tes. Analisis tes adalah salah satu kegiatan dalam rangka mengkonstruksi tes untuk mendapatkan gambaran tentang mutu tes, baik mutu keseluruhan tes maupun mutu tiap butir soal/tugas. Analisis dilakukan setelah tes disusun dan dicobakan kepada sejumlah subyek dan hasilnya menjadi umpan balik untuk perbaikan/peningkatan mutu tes bersangkutan. Oleh karena itu kegiatan analisis tes merupakan keharusan dalam keseluruhan proses mengkonstruksi tes. Kegiatan analisis tes mencakup empat hal diantaranya 1) analisis validitas tes, 2) analisis reliabilitas tes, 3) analisis butir soal dan 4) analisis teknis kegunaan soal.

Kegiatan analisis butir soal merupakan kegiatan penting dalam penyusunan soal agar diperoleh butir soal yang bermutu. Analisis butir soal yang paling sederhana yang dapat dilakukan adalah analisis tingkat kesukaran butir soal dan analisis daya pembeda butir soal.

Tujuan kegiatan analisis butir soal adalah mengkaji dan menelaah setiap butir soal agar diperoleh soal yang bermutu sebelum digunakan, meningkatkan kualitas butir tes melalui revisi atau membuang soal yang tidak efektif, serta mengetahui informasi diagnostik pada mahasiswa apakah mereka telah memahami materi yang telah diajarkan. Soal yang bermutu adalah soal yang dapat memberikan informasi setepat-tepatnya tentang mahasiswa yang telah menguasai materi dan yang belum menguasai materi (Suprananto, 2012).

Fakta di lapangan menunjukkan bahwa analisis butir soal masih jarang dilakukan oleh pendidik. Beberapa

alasan yang menyebabkan pendidik tidak melakukan analisis butir soal antara lain: (1) pendidik merasa terbebani dalam proses analisis butir soal sehingga tidak melakukannya, (2) pendidik meyakini bahwa kualitas soal tes yang dibuat sudah baik sehingga tidak melakukan penelaahan lebih lanjut. Berdasarkan uraian di atas, maka perlu dilakukan analisis butir soal secara kuantitatif yang untuk mengetahui kualitas dan reliabilitas soal tersebut. Kualitas soal kompetensi pada penelitian ini dapat dilihat dari hasil uji tingkat kesukaran, daya beda, dan validitas butir soal. Sebenarnya banyak pendapat ahli yang mendefinisikan makna dari tes itu. Brown (1976) dalam Susanto (2015) mengatakan bahwa tes adalah prosedur yang sistematis guna mengukur sampel perilaku seseorang. Menurut Cronbach dalam bukunya yang berjudul *Essential of Psychological Testing* yang dikutip oleh Susanto (2015) menyatakan bahwa tes adalah suatu prosedur yang sistematis untuk membandingkan tingkah laku dua orang atau lebih. Suryabrata (2003) mengatakan tes adalah pertanyaan-pertanyaan yang harus dijawab dan atau perintah-perintah yang harus dijalankan, yang mendasarkan harus bagaimana *testee* menjawab pertanyaan-pertanyaan atau melakukan perintah-perintah itu, penyelidik mengambil kesimpulan dengan cara membandingkan dengan standart atau *testee* lainnya. Menurut Buchori dalam Arikunto mengatakan bahwa tes adalah suatu percobaan yang diadakan untuk mengetahui ada atau tidaknya hasil-hasil pelajaran tertentu pada seorang murid atau kelompok murid. Dalam bukunya yang berjudul *Evaluasi Pendidikan*, Arikunto juga mengatakan bahwa tes adalah suatu alat atau prosedur yang sistematis dan obyektif untuk memperoleh data-data atau keterangan-keterangan yang

diinginkan tentang seseorang, dengan cara yang boleh dikatakan tepat dan cepat. Dari beberapa uraian di atas dapat disimpulkan bahwa tes adalah prosedur yang sistematis, obyektif dan standart yang berupa serentetan pertanyaan atau latihan yang harus dijawab oleh *testee* untuk menghasilkan suatu nilai yang mencerminkan tingkah laku atau prestasi *testee*.

Secara umum, ada dua macam fungsi yang dimiliki oleh tes, yaitu 1) sebagai alat pengukur terhadap peserta didik. Dalam hubungan ini tes berfungsi mengukur tingkat perkembangan atau kemajuan yang telah dicapai oleh peserta didik setelah mereka menempuh proses belajar- mengajar dalam jangka waktu tertentu, 2) sebagai alat pengukur keberhasilan program pengajaran, sebab melalui tes tersebut akan dapat diketahui sudah seberapa jauh program pengajaran yang telah ditentukan, telah dapat dicapai. Secara umum tes dibedakan berdasarkan obyek pengukurannya dapat dibagi menjadi dua, yaitu tes kepribadian (*personality test*) dan tes hasil belajar (*Achievement test*). Sebuah tes dikatakan baik apabila memenuhi syarat syarat tes diantaranya 1) Validitas, 2) Reliabilitas, 3) Obyektivitas, 4) Praktibilitas (*Practibility*) dan 5) Ekonomis.

Salah satu bentuk evaluasi yang menggunakan tes adalah evaluasi penilaian pencapaian kompetensi pada mahasiswa kebidanan. Bentuk evaluasi seperti tertuang pada kurikulum D3 Kebidanan bertujuan menilai pencapaian standar kompetensi lulusan untuk semua mata kuliah dalam bentuk Penilaian Pencapaian Kompetensi (PPK) I, II dan III yang terjabarkan dalam beberapa tahap, yaitu 1) Penilaian Pencapaian Kompetensi (PPK) I meliputi kompetensi penguasaan ilmu-ilmu dasar dan keterampilan dasar yang terkait dengan

praktik kebidanan, 2) Penilaian Pencapaian Kompetensi (PPK) II meliputi kompetensi asuhan kebidanan yang meliputi kehamilan, persalinan, dan bayi baru lahir, nifas, neonatus, bayi, balita dan pra sekolah, komunitas, kesehatan reproduksi dan keluarga berencana serta kegawatdaruratan maternal dan neonatal, 3) Penilaian Pencapaian Kompetensi (PPK) III meliputi kompetensi komprehensif melalui uji penampilan klinik sebagai kandidat bidan dan laporan tugas akhir. Evaluasi PPK ini dilaksanakan secara bertahap dan berkesinambungan setelah pembelajaran teori telah selesai.

Merujuk pada latar belakang diatas dan mengingat pentingnya analisis tes ini dilakukan maka peneliti tertarik untuk melakukan penelitian ini dengan tujuan untuk mengetahui hasil analisis tes penilaian pencapaian kompetensi (PPK) ditinjau dari tingkat kesukaran, daya beda, validitas butir dan reliabilitas tes.

METODE

Penelitian ini merupakan jenis penelitian deskriptif kuantitatif dengan pendekatan analitik. Sampel yang dimaksud pada penelitian ini adalah perangkat tes penilaian pencapaian kompetensi yang terdiri dari 120 butir soal pilihan ganda dengan menggunakan skala dikotomi. Subjek penelitiannya adalah seluruh mahasiswa Jurusan Kebidanan semester V yang mengikuti Ujian Penilaian Pencapaian Kompetensi (PPK) Tahap I. Kualitas butir tes PPK ini meliputi tingkat kesukaran, daya beda dan validitas butir soal dan reliabilitas tes.

Bilangan yang menunjukkan sukar dan mudahnya sesuatu soal disebut dengan indeks kesukaran. Besarnya

indeks kesukaran dilihat pada tabel berikut ini (Koyan, W.2007).

Tabel 1. Interpretasi Tingkat Kesukaran Soal

p	Interpretasi
0.0 - 0.30	: Soal sukar
0.31 - 0.70	: Soal sedang
0.71 - 1.0	: Soal mudah

Rumus mencari indeks kesukaran menurut Daryanto (2009) adalah :

$$P = \frac{B}{JS} \quad (1)$$

Dimana :

P = indeks kesukaran

B = banyaknya siswa yang menjawab soal itu dengan betul

JS = jumlah seluruh siswa peserta tes.

Daya pembeda soal yaitu kemampuan sesuatu soal untuk membedakan antara siswa yang berkemampuan tinggi dengan siswa yang berkemampuan rendah. Angka yang menunjukkan besarnya daya pembeda disebut indeks diskriminasi (D), dan nilainya berkisar antara 0,00 sampai 1,00. Pada daya pembeda ini berlaku tanda negatif yang digunakan jika sesuatu soal “terbalik” menunjukkan kualitas testee yaitu anak pandai disebut bodoh dan anak bodoh disebut pandai. Bagi suatu soal yang dapat dijawab benar oleh siswa kemampuan tinggi dan siswa kemampuan rendah, maka soal itu tidak baik karena tidak punya daya pembeda. Demikian juga jika semua kelompok bawah menjawab salah dan siswa berkemampuan tinggi juga sama-sama menjawab salah, maka soal itu tidak mempunyai daya beda sama sekali.

Cara menentukan daya pembeda (nilai D) yaitu perlu dibedakan antara kelompok kecil (kurang dari 100) dan kelompok besar (100 orang ke atas). Mengingat biaya dan waktu menganalisis, maka untuk kelompok besar biasanya hanya diambil dua kutub

saja yaitu 27% skor teratas sebagai kelompok atas (JA) dan 27 % skor terbawah sebagai kelompok bawah (JB), sedangkan untuk kelompok kecil seluruh kelompok tes dibagi dua sama besar, 50% kelompok atas dan 50% kelompok bawah.

Rumus mencari daya pembeda menurut Daryanto (2009) yaitu :

$$DP = \frac{B_A}{J_A} - \frac{B_B}{J_B} = P_A - P_B \quad (2)$$

Dimana :

D : Daya Pembeda

J : Jumlah peserta tes

JA : Jumlah peserta kelompok atas

JB : Jumlah peserta kelompok bawah

BA : banyaknya peserta kelompok atas yang menjawab soal dengan benar

BB : banyaknya peserta kelompok bawah yang menjawab soal dengan benar

PA : proporsi peserta kelompok atas yang menjawab benar (ingat P sebagai indeks kesukaran)

PB : Proporsi peserta kelompok bawah yang menjawab benar

Interpretasi daya pembeda dapat dilihat pada tabel berikut.

Tabel 2. Interpretasi Daya Pembeda Soal

Indek Diskriminasi	Interpretasi
0.0 - 0.2	: jelek
0.2 - 0.4	: cukup
0.4 - 0.7	: baik
0.7 - 1.0	: Baik sekali

Nilai D yang negatif artinya soal tersebut tidak baik, jadi semua butir soal yang mempunyai nilai D negatif sebaiknya dibuang saja.

Validitas adalah suatu ukuran yang menunjukkan tingkat kesahihan suatu tes. Suatu tes dikatakan valid apabila tes tersebut mengukur apa yang hendak diukur. Tes memiliki validitas yang tinggi jika hasilnya sesuai dengan kriteria, dalam arti memiliki kesejajaran antara tes dan kriteria (Arikunto, S.2006). Untuk menguji validitas setiap butir soal dapat dilakukan dengan uji validitar butir tes dikotomi, benar mendapat skor 1 dan salah akan mendapat skor 0. Untuk menguji validitas instrumen tes digunakan *rumus point biserial* yaitu (Sugiyono, 2009):

$$r_{pbi} = \frac{M_{pi} - M_t}{St} \sqrt{\frac{p_i}{q_i}} \quad (3)$$

Keterangan:

- rp : Koefisien korelasi biserial
- bi
- M : Rerata skor subjek
- p : menjawab benar pada butir ke-i
- Mt : Rerata skor total
- St : Standar deviasi skor total
- pi ; Peluang menjawab benar butir ke-i
- qi : Peluang menjawab salah butir ke-i

Untuk menginterpretasikan tingkat validitas, maka dapat dilakukan dengan membandingkan nilai r hitung dengan r kritis tabel biserial. Jika nilai r hitung lebih besar dari r kritis tabel maka Ho ditolak yang artinya soal tersebut valid.

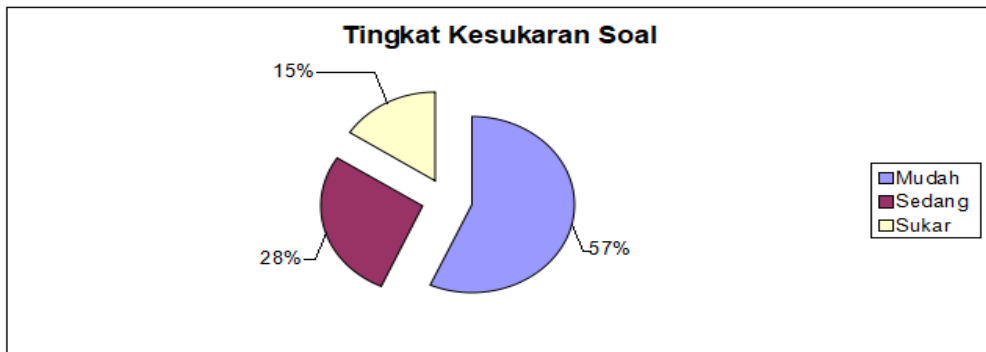
Sedangkan tes reliabilitas instrumen adalah keadaan instrumen yang menunjukkan hasil pengukuran yang reliable (tidak berubah-ubah, konsisten). Instrumen yang reliable adalah instrumen yang apabila digunakan untuk mengukur subyek atau objek yang sama pada waktu yang berbeda dan pengukuran dilakukan oleh orang yang berbeda hasilnya tetap sama.

Beberapa faktor penting yang mempengaruhi reliabilitas suatu tes yaitu kemampuan peserta tes atau subjek uji coba, 2) jumlah peserta tes, 3) panjang pendeknya tes., 4) evaluasi yang subjektif juga akan menurunkan reliabilitas. Reliabilitas instrumen dinyatakan dengan koefisien reliabilitas. Instrumen yang reliable adalah instrumen yang memiliki koefisien reliabilitas minimal 0,70. Sebaiknya koefisien reliabilitas instrumen 0,80 atau lebih. Koefisien reliabilitas instrumen dihitung dengan menggunakan rumus Kuder Richardson-20 (KR-20).

HASIL DAN PEMBAHASAN

Fokus penelitian ini adalah analisis butir soal penilaian pencapaian kompetensi pada mahasiswa kebidanan.

Hasil analisis tingkat kesukaran soal oleh gambar dibawah ini. merujuk pada rumus no (1) ditunjukkan



Gambar 1. Diagram Pie Hasil Analisis Tingkat Kesukaran Butir Soal Penilaian Pencapaian Kompetensi pada Mahasiswa Kebidanan Tahun 2018

Pada gambar diatas terlihat dari 120 butir soal didapatkan 68 (57%) butir soal menunjukkan tingkat kesukaran mudah, 33 (28%) butir soal menunjukkan tingkat kesukaran sedang dan 19 (16%) butir soal menunjukkan tingkat kesukaran sukar. Hal tersebut menunjukkan proporsi soal pada perangkat tes PPK tersebut masih belum seimbang. Keseimbangan yang dimaksud adalah jumlah soal-soal yang tergolong mudah, sedang dan sukar seimbang. Tingkat kesukaran soal perlu dilihat dari kemampuan mahasiswa dalam menjawab soal yang diberikan, bukan dilihat dari sudut pandang dosen sebagai pembuat soal. Perbandingan antara soal yang mudah-sedang-sukar dapat dibuat 3-4-3 atau 3-5-2, yang diartinya adalah 30% soal kategori mudah, 40% soal kategori sedang, dan 30% soal kategori sukar; atau 30% soal kategori mudah, 50% soal kategori sedang dan 20% soal kategori sukar. Hasil penelitian menunjukkan bahwa 57% soal pencapaian penilaian kompetensi terkategori mudah, namun pada kenyataannya masih banyak yaitu 37 orang dari 78 orang (47%) mahasiswa yang belum dinyatakan lulus tes. Hal ini menunjukkan bahwa penguasaan materi oleh mahasiswa secara konseptual terkait

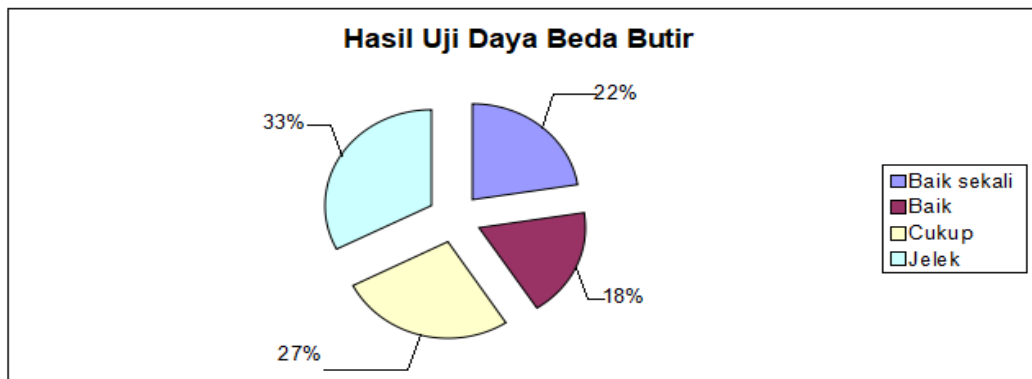
dengan kompetensi bidang masih rendah. Disisi lain hal tersebut bisa terjadi karena variasi jumlah soal yang cukup banyak sehingga memungkinkan mahasiswa untuk tidak dapat menjawab pada komponen soal yang berbeda. Pertimbanganyang lain adalah jenis soal, sesuai dengan tuntutan kurikulum D3 Kebidanan yang menyatakan bahwa soal yang dibuat harus bermuatan kompetensi, yaitu berupa soal kasus. Hal ini tentu menyebabkan dari segi isi (*content*) soal penilaian kompetensi ini cenderung lebih sulit karena memerlukan proses analisa dan waktu yang lebih lama.

Butir-butir soal yang masuk dalam kategori sedang sebaiknya segera dicatat dalam buku bank soal. Selanjutnya, butir-butir soal tersebut dapat dikeluarkan lagi dalam tes pada waktu yang akan datang. Untuk butir soal yang sukar ada tiga kemungkinan tindak lanjut, yaitu: (1) butir soal tersebut dibuang dan tidak dikeluarkan lagi dalam tes yang akan datang, (2) diteliti ulang, dilacak dan ditelusuri sehingga dapat diketahui faktor apa yang menyebabkan butir item tersebut sulit dijawab oleh mahasiswa. Setelah dilakukan perbaikan, butir-butir soal tersebut dikeluarkan lagi dalam tes hasil

belajar yang akan datang, (3) butir soal yang sukar sebaiknya digunakan, dalam tes seleksi yang sifatnya ketat karena kondisi tersebut mendukung penggunaan soal yang sukar dengan asumsi bahwa mahasiswa yang berkemampuan cukup memadai akan lolos dalam seleksi yang diadakan. Butir soal yang mudah juga memiliki tiga kemungkinan tindak lanjut yaitu: (1) Butir soal tersebut dibuang dan tidak dikeluarkan lagi dalam tes yang akan datang, (2) diteliti ulang dan ditelusuri secara cermat guna mengetahui faktor yang menyebabkan butir soal tersebut dapat dijawab betul oleh seluruh mahasiswa. Setelah dilakukan perbaikan butir soal yang bersangkutan dikeluarkan

lagi pada tes berikutnya untuk mengetahui apakah derajat kesukaran butir soal menjadi lebih baik daripada sebelumnya atau tidak, (3) butir-butir soal yang mudah mempunyai manfaat yaitu butir-butir soal dapat dimanfaatkan pada tes seleksi yang bersifat longgar, dalam arti bahwa sebagian besar dari mahasiswa akan dinyatakan lulus dalam tes seleksi tersebut. Dalam kondisi ini, pemberian butir soal yang mudah akan memberikan kesempatan bagi banyak mahasiswa untuk lolos dalam tes seleksi atau ujian yang diadakan.

Dengan menggunakan rumus (2) diperoleh hasil daya pembeda butir soal seperti pada gambar berikut ini.



Gambar 2. Diagram Pie Hasil Analisis Daya Beda Butir Soal Penilaian Pencapaian Kompetensi pada Mahasiswa Kebidanan Tahun 2018

Berdasarkan analisis yang peneliti lakukan diperoleh informasi bahwa dari 120 butir soal tersebut terdapat 39 (33%) butir soal yang daya pembedanya terkategori jelek, 32 (27%) butir soal terkategori cukup, 22 (18%) butir soal terkategori baik dan 27 (22%) terkategori baik sekali. Adapun tindak lanjut atas hasil analisis mengenai daya pembeda item tes tersebut adalah soal yang memiliki daya pembeda yang baik hendaknya dimasukkan dan dicatat dalam buku bank soal. Butir soal tersebut dapat dikeluarkan lagi dalam tes selanjutnya karena kualitasnya sudah

cukup memadai. Butir soal yang daya pembedanya masih jelek memiliki dua kemungkinan tindak lanjut yaitu: (1) ditelusuri dan diperbaiki sehingga nanti dapat diajukan lagi dalam tes yang akan datang, serta perlu dianalisis kembali apakah daya pembeda meningkat atau tidak, (2) dibuang dan tidak dikeluarkan kembali pada tes yang akan datang. Khusus butir soal yang angka indeks diskriminanya bertanda negatif sebaiknya tidak perlu dikeluarkan pada tes hasil belajar selanjutnya atau direvisi kembali.

Hasil analisis validitas soal

menunjukkan bahwa dari 120 butir soal yang diuji validitasnya dengan menggunakan uji poin biserial, didapatkan hanya 43 (36%) butir soal yang dinyatakan valid. Hal ini menunjukkan bahwa butir soal penilaian pencapaian kompetensi yang terkategori tidak valid (drop) harus direvisi kembali atau ditinjau ulang jika ingin digunakan lagi pada tes berikutnya, dan indeks validitas yang memiliki nilai negatif harus dibuang dan tidak dapat digunakan kembali pada tes berikutnya. Hal ini dapat terjadi karena perangkat tes penilaian pencapaian kompetensi yang digunakan selama ini hanya dianalisis secara kualitatif yakni melalui review soal yang dilakukan oleh dosen pengampu yang expert di bidangnya. Padahal disisi lain analisis secara kuantitatif uji validitas butir soal ini belum pernah dilakukan. Hal inilah yang menyebabkan banyak soal yang dinyatakan tidak valid (drop).

Uji reliabilitas pada perangkat tes ini dilakukan dengan membuang terlebih dahulu butir soal yang dinyatakan tidak valid atau dengan kata lain pengujian dilakukan terhadap butir soal yang valid saja. Dari uji reliabilitas pada 43 butir soal yang valid didapatkan nilai reliabilitas 0.95 yang artinya reliabilitas soal tersebut sangat baik. Pengujian reliabilitas pada perangkat tes penilaian kompetensi ini dilakukan dengan menggunakan formula KR 20 yang diperuntukkan untuk menghitung reliabilitas soal yang terkategori dikotomi. Reliabilitas mengandung pengertian derajat konsistensi atau keterandalan dari suatu perangkat tes, artinya jika suatu tes dilakukan berulang-ulang akan menunjukkan hasil yang sama. Melihat koefisien reliabilitas pada penelitian ini sudah sangat baik hal ini menunjukkan bahwa butir-butir soal yang ada didalam perangkat tes tersebut dapat digunakan pada situasi dan sampel

yang berbeda.

SIMPULAN DAN SARAN

Hasil penelitian mengenai analisis tes penilaian pencapaian kompetensi dapat disimpulkan 1) tingkat kesukaran butir soal pada tes penilaian pencapaian kompetensi sebagian besar berada pada rentang mudah, 2) daya pembeda butir soal pada tes penilaian kompetensi sebagian besar berada pada kategori jelek, 3) uji validitas butir soal pada tes penilaian pencapaian kompetensi menunjukkan sebagian besar butir soal masih belum valid, dan 4) tingkat reliabilitas soal pada butir soal yang valid pada tes penilaian pencapaian kompetensi menunjukkan tingkat reliabilitas yang sangat baik.

Merujuk pada hasil penelitian ini penulis berharap lembaga lebih memfasilitasi dosen dalam melakukan uji validitas dan reliabilitas soal dengan menambah sumber daya manusia yang dapat membantu meringankan tugas dosen dalam melakukan analisis butir soal. Kepada dosen diharapkan lebih memahami akan pentingnya alat ukur evaluasi, karena salah satu keberhasilan pembelajaran ditentukan oleh baik tidaknya alat ukurnya. Tes sebagai salah satu alat ukur evaluasi hendaknya dapat dianalisis secara kuantitatif meliputi tingkat kesukaran, daya beda, validitas dan reliabilitas tes.

DAFTAR PUSTAKA

- Arikunto, S. 2006. *Prosedur Penelitian Suatu Pendekatan Praktik*. Edisi revisi VI. PT Rineka Cipta. Jakarta
- Arikunto, S. 2009. *Dasar-dasar Evaluasi Pendidikan*. Edisi revisi. PT Bumi Aksara. Jakarta
- Daryanto. 2009. *Panduan Proses Pembelajaran Kreatif dan Inovatif*. AV Publisher. Jakarta

- Depdiknas. 2007. *Analisis Butir Soal Secara Manual*. Pusat Penilaian Pendidikan, Badan Penelitian dan Pengembangan. Depdiknas. Bogor
- Kerlinger, Fred N. 1990. *Asas-Asas Penelitian Behavioral*. Terjemahan Landung R.Simatupang. *Foundation of Behavioral Research*. 1973. Yogyakarta. Gadjah Mada University Press
- Koyan,W.2007. *Statistik Terapan (Teknik Analisis Data Kuantitatif)*.Buku Ajar.Universitas Pendidikan Ganesha.Singaraja
- Koswara, Triatna. 2009. *Manajemen Peningkatan Mutu Pendidikan*. Cetakan Ketiga. Alfabeta. Bandung
- Kemenristek RI. 2012. *Pedoman Penilaian Pencapaian Kompetensi Program Diploma III Kebidanan*. Jakarta
- Nazir,M. 2009. *Metode Penelitian*.Cetakan Ketujuh. Penerbit Ghalia Indonesia.
- Sudijono, Anas. 2013. *Pengantar Evaluasi Pendidikan*. Penerbit Rajawali Press. Jakarta
- Sugiyono.2012. *Statistika Untuk Penelitian*. Cetakan Kelima. CV Alfabeta. Jakarta
- Sukardi. 2011. *Evaluasi Pendidikan: Prinsip dan Operasionalnya*. Bumi Aksara. Jakarta
- Suprananto, Kusaeri. 2012.*Pengukuran dan Penilaian Pendidikan*. Graha Ilmu. Yogyakarta
- Suryabrata,S.2003. *Metodologi Penelitian*. PT Raja Grafindo Persada. Jakarta
- Susanto, Hery. 2015. Analisis Validitas Reliabilitas Tingkat Kesukaran dan Daya Beda Pada Butir Soal Ujian Akhir Semester Ganjil Kelas XII IPS Di SMA Negeri 12 Bandar Lampung Tahun Ajaran 2014/2015. *Jurnal Ilmu Pendidikan*. Lampung