

**THE DEVELOPMENT OF COGNITIVE TEST DEVICES FOR  
COMPETENCE MEASUREMENT OF FUTURE SENIOR VOCATIONAL  
SCHOOL TEACHERS IN FISHING VESSEL NAUTICAL EXPERTISE**

**PENGEMBANGAN PERANGKAT TES KOGNITIF  
UNTUK PENGUKURAN KOMPETENSI CALON GURU  
SMK KEAHLIAN NAUTIKA KAPAL PENANGKAP IKAN**

Oleh :

Eliza Merina dan Budi Susetyo  
PPPPTK Pertanian Cianjur  
Email : elizamerina@gmail.com

**Abstract.** *This research aimed to develop a valid and reliable cognitive test device for measuring the professional competence of the future teachers in the fishing vessel nautical science with a proper procedure. The method used in this study was research and development method with the following procedures: identification of measurement purposes, the development of test specifications, item writing, item review of the test (validity test), trial test 1, items analysis 1, reliability test, trial test 2, items analysis 2, and the construction of the final test device. At the beginning of the development of test specifications and test item writing, there were 65 test items being developed and as the development going, the number of test items was decreasing. After being tested by 3 experts in the fishing vessel nautical science, the items have been reduced to 60 test items which then were being constructed to be the 1<sup>st</sup> trial test device. Following this, the 1<sup>st</sup> items analysis was carried out in accordance to classical test theory which consists the level of difficulties, item discrimination, and effectiveness of distractors. From this analysis the number of items has been reduced to 29 test items which were constructed to be 2<sup>nd</sup> trial test device. Reliability test was conducted for the 1<sup>st</sup> trial test using the internal consistency approach by calculating the reliability of Kuder Richardson ( $KR_{20}$ ). The result showed that the reliability coefficient is 0.71 which indicated that the 1<sup>st</sup> trial test was a reliable test. The 2<sup>nd</sup> trial test was conducted and followed by the 2<sup>nd</sup> items analysis which left 19 test items to be constructed as a final form test devices.*

**Keywords** : *test development, competency measurement, fishing vessel nautical science*

**Abstrak.** Penelitian ini dilakukan dengan tujuan untuk mengembangkan sebuah perangkat tes kognitif yang valid dan reliabel untuk pengukuran kompetensi profesional calon-calon guru SMK Keahlian NKPI melalui prosedur yang benar. Metode penelitian yang digunakan dalam penelitian ini adalah riset dan pengembangan. Prosedur yang dilalui dalam penelitian ini yaitu : identifikasi tujuan pengukuran, pengembangan spesifikasi tes, penulisan butir soal, penelaahan butir tes (uji validitas), uji coba tes 1, analisa butir tes 1, uji reliabilitas, uji coba tes 2, analisa butir tes 2, dan perakitan bentuk akhir tes. Pada awal pembuatan spesifikasi tes dan penulisan butir, jumlah butir yang dikembangkan berjumlah 65 butir tes. Seiring proses pengembangan tes dilakukan, jumlah butir tes mengalami degradasi (penurunan). Setelah melalui tahap telaah butir (uji validitas) yang dilakukan oleh 3 orang ahli di bidang NKPI, butir tes mengalami pengurangan 5 butir tes sehingga tersisa 60 butir tes untuk dirakit menjadi perangkat tes uji coba 1. Setelah uji coba 1 dilakukan dilanjutkan dengan analisis butir 1 yang dilakukan sesuai teori ujian klasik yang terdiri dari tingkat kesukaran, daya diskriminasi butir, dan efektivitas distraktor. Dari analisis tersebut, jumlah butir soal pun mengalami pengurangan hingga tersisa 29 butir tes untuk dirakit menjadi perangkat tes uji coba 2. Dari hasil uji coba 1 pun dilakukan uji reliabilitas menggunakan pendekatan konsistensi internal (*internal consistency*) yaitu dengan perhitungan reliabilitas Kuder Richardson ( $KR_{20}$ ). Diperoleh nilai koefisien reliabilitas sebesar 0,71 yang menunjukkan bahwa perangkat tes tergolong reliabel. Uji coba 2

dilakukan dan diteruskan dengan analisa butir 2, kembali lagi butir soal mengalami pengurangan hingga tersisa 19 butir tes untuk dirakit menjadi perangkat tes bentuk akhir.

**Kata Kunci** : pengembangan tes, pengukuran kompetensi, nautika kapal penangkapan ikan

## A. PENDAHULUAN

Peraturan Pemerintah Nomor 74 Tahun 2008 Tentang Guru menerangkan bahwa guru harus memiliki empat kompetensi dalam melaksanakan tugas keprofesionalannya. Untuk itu diperlukan upaya-upaya untuk memastikan hal tersebut, diantaranya dengan adanya seleksi calon guru. Dengan demikian adanya perangkat tes sebagai alat ukur untuk menyeleksi sangat diperlukan. Penelitian ini dilakukan dengan tujuan untuk mengembangkan sebuah perangkat tes kognitif yang valid dan reliabel untuk pengukuran kompetensi profesional calon-calon guru SMK Keahlian NKPI melalui prosedur yang benar. Perangkat tes yang akan dikembangkan dalam penelitian ini tidak dimaksudkan untuk mengukur seluruh kompetensi tersebut, namun khusus kompetensi profesional dari calon-calon guru NKPI saja. Dengan demikian, perangkat tes yang akan dikembangkan dalam penelitian ini dimaksudkan untuk mengukur sejauhmana penguasaan para calon guru NKPI secara kognitif terhadap ilmu pengetahuan dan teknologi dalam bidang NKPI. Pokok permasalahan dalam penelitian ini adalah bagaimanakah prosedur penyusunan perangkat tes kognitif terstandar bagi calon guru SMK

Keahlian NKPI sehingga diperoleh perangkat tes yang valid dan reliabel. Hasil penelitian ini dapat dimanfaatkan untuk kegiatan seleksi guru guna berbagai tujuan seperti seleksi calon guru dalam PPG, sebagai dasar pengambilan keputusan bagi pimpinan di sekolah, pemilihan guru teladan, dan sebagainya.

## B. PEMBAHASAN

Tes dapat didefinisikan sebagai suatu pertanyaan atau tugas atau seperangkat tugas yang direncanakan untuk memperoleh informasi tentang trait atau atribut pendidikan atau psikologik yang setiap butir pertanyaan atau tugas tersebut mempunyai jawaban atau ketentuan yang dianggap benar. (Zainul dan Nasoetion, 1993 : 2). Definisi lain dapat dikutip dari pendapat Lee J Cronbach (Azwar, 2013 : 3) dalam bukunya *Essentials of Psychological Testing* yang menjelaskan bahwa tes adalah “... a systematic procedure for observing a person's behavior and describing it with the aid of a numerical scale or a category system”.

Susetyo (2011 : 39 – 54) menjelaskan bahwa pembuatan perangkat tes dilakukan dalam dua kegiatan yang saling bergantung satu dengan yang lainnya. Kegiatan pertama yaitu mendesain fisik perangkat tes atau perangkat ukur.

Kegiatan semacam ini dilakukan oleh ahli pengukuran atau psikometri. Kegiatan yang kedua adalah pembuat tes, yaitu orang yang memiliki keahlian dalam isi atau materi tes yang akan dibuat.

Metode penelitian yang digunakan dalam penelitian ini adalah riset dan pengembangan, sebab dari penelitian ini tidak menutup kemungkinan dilakukannya penelitian lanjutan baik untuk perbaikan perangkat tes yang telah tersusun ataupun untuk pengembangan perangkat tes lainnya. Ali (2011 : 394) menjelaskan bahwa “riset dan pengembangan (*research and development*) pada hakikatnya merupakan suatu proses dalam mengembangkan dan memvalidasi perangkat tertentu yang menjadi produknya ...” Metode penelitian ini sering juga disebut dengan pengembangan berbasis riset.

Berikut adalah proses pengembangan perangkat tes beserta hasilnya dari penelitian yang penulis lakukan :

### **1. Identifikasi Tujuan Pengukuran**

Identifikasi tujuan merupakan penegasan tujuan pengukuran yang akan dicapai oleh tes yang diikuti oleh pembatasan kawasan ukur, yakni pendefinisian lingkup materi ukur yang hendak diungkap (Azwar, 2013 : 55). Perangkat tes kognitif ini disusun untuk mengukur kompetensi

profesional calon-calon guru SMK Keahlian NKPI. Perangkat tes yang akan dikembangkan dalam penelitian ini dimaksudkan untuk mengukur kompetensi profesional saja sehingga tes yang disusun hanya akan berbicara seputar pengetahuan bidang ilmu pengetahuan dan teknologi terkait bidang Nautika Kapal Penangkap Ikan (NKPI).

### **2. Pengembangan Spesifikasi Tes**

Dari tujuan pengukuran yang spesifik dan konkret, dapat disusun suatu tabel spesifikasi. Tabel ini merupakan suatu pedoman yang akan menjaga agar penulisan butir tetap terarah pada tujuan pengukuran tes dan tidak keluar dari batasan isi (Azwar, 2013 : 69). Tabel spesifikasi yang disusun dalam penelitian ini mengacu pada kurikulum SMK 2013, KKNI, dan kurikulum pendidikan sarjana.

Struktur Kurikulum SMK Paket Keahlian NKPI digunakan sebagai dasar penyusunan spesifikasi tes. Mengingat perangkat tes yang dikembangkan dalam penelitian ini ditujukan untuk mengukur kompetensi profesional calon-calon guru SMK Paket Keahlian NKPI dan dengan memperhatikan KKNI serta kurikulum pendidikan sarjana (S1) Program Studi Teknologi dan

Manajemen Perikanan Tangkap (TMPT), maka hanya mata pelajaran dari kelompok C2 dan C3 saja yang digunakan. Sehingga ada sembilan mata pelajaran yang dijadikan dimensi dalam pengembangan spesifikasi tes yaitu :

- 1) Hukum Maritim, Peraturan Perikanan/*CCRF* dan Pencegahan Polusi Lingkungan Laut (HMP4L2)
- 2) Bangunan dan Stabilitas Kapal Perikanan (BSKP)
- 3) Dasar-dasar Teknik Penangkapan Ikan, Penanganan dan Penyimpanan Hasil Tangkap (DTPIP2HT)
- 4) Dasar-Dasar Keselamatan di Laut (DKL)
- 5) Bahasa Inggris Maritim dan Perikanan (BIMP)
- 6) Pelayaran Kapal Perikanan (PKP)
- 7) Komunikasi dan Dinas Jaga/P2TL Kapal Perikanan (KDJKP)
- 8) Manajemen Kapal Perikanan (MKP)
- 9) Teknik dan Alat Penangkap Ikan (TAPI)

Spesifikasi tes disusun ke dalam sebuah tabel dimana mata pelajaran berfungsi sebagai dimensi tes. Dari setiap mata pelajaran ini dijabarkan

Kompetensi Inti (KI) yang melingkupinya. Kompetensi Inti (sebelumnya adalah Standar Kompetensi (SK)) merupakan kemampuan dasar yang merujuk kepada klasifikasi kemampuan minimal peserta didik pada aspek-aspek kemampuan dalam pengetahuan, afektif, dan keterampilan dari suatu program pendidikan tertentu (Susetyo, 2011 : 75). Dari Kompetensi Inti (KI) dijabarkan pula Kompetensi Dasar (KD) beserta indikator ketercapaiannya. Kompetensi Dasar (KD) merupakan penjabaran dari Kompetensi Inti (KI), berisikan deskripsi dari isi tujuan yang terkandung di dalamnya dan sebagai acuan pencapaian tujuan kegiatan pembelajaran dari program pendidikan tertentu (Susetyo, 2011 : 76). Indikator merupakan penjabaran dari Kompetensi Dasar (KD), berkaitan dengan topik pembahasan (materi) dari suatu program pembelajaran tertentu (Susetyo, 2011 : 76). Baik KI maupun KD, keduanya diambil dari Kurikulum SMK 2013.

Dengan merujuk kepada KI, KD, dan indikator ketercapaian maka dibuat estimasi tingkat kesukaran butir soal, dimensi kognitif soal, dan jumlah soal dari setiap dimensinya.

Untuk melihat gambaran kemampuan peserta tes secara lengkap maka butir-butir tes yang dibuat pada tingkat kesukaran mudah, sedang, dan sukar. Dengan mempertimbangkan bahwa perangkat tes ini ditujukan bagi calon-calon guru, maka setiap butir tes dibuat pada tingkatan kognitif kedua, ketiga, dan keempat yaitu pemahaman (C2), aplikasi (C3), dan analisis (C4). Mengenai jumlah butir tes, Susetyo (2011 : 84) menyebutkan bahwa jumlah butir tes yang terdapat dalam suatu perangkat tes tidak dapat ditetapkan secara umum. Azwar (2013 : 77 – 78) memaparkan bahwa menentukan banyaknya butir menyangkut pertimbangan teoritik maupun praktis. Secara teoritik, suatu tes haruslah berisi sebanyak-banyaknya butir yang independen (tidak terikat) satu sama lain. Namun secara praktis, harus pula dipertimbangkan masalah-masalah seperti tujuan diadakan tes, waktu yang tersedia bagi penulisan butir dan pemeriksaan jawaban tes, jumlah peserta yang akan dikenai tes, waktu yang tersedia bagi peserta untuk menjawab tes, dan sebagainya. Dengan demikian, dengan mempertimbangkan hal-hal tersebut jumlah 65 butir tes dirasakan cukup untuk penelitian ini.

Pada tahapan ini pula ditentukan bentuk tes yang akan dibuat juga lama waktu pengerjaan tes. Bentuk soal yang dirasakan sesuai untuk mengukur penguasaan calon-calon guru NKPI terhadap kompetensi bidang NKPI adalah pilihan ganda. Mengapa pilihan ganda? Kompetensi dalam bidang NKPI bukanlah hal yang sederhana. Sehingga diperlukan bentuk tes yang komprehensif dan mampu memuat banyak butir dalam waktu yang singkat. Bentuk tes pilihan ganda dapat memberikan kemudahan pula dalam pemeriksaan jawaban dan pemberian skor secara cepat namun tetap terjaga objektivitasnya meskipun jumlah peserta tes banyak. Selain itu, pada bentuk tes pilihan ganda kualitas butir tes dapat dianalisis secara empirik sehingga menjadikannya tidak mudah untuk dalam proses pembuatannya (Azwar, 2013 : 75). Setiap butir soal dapat dikerjakan antara 15 detik hingga 60 detik, tergantung tingkat kesukaran setiap soal. Oleh karena itu, dengan mempertimbangkan jumlah dan tingkat kesukaran butir tes maka diperkirakan waktu yang cukup untuk menjawab 65 butir soal tersebut adalah 30 menit.

### **3. Penulisan Butir Soal**

Penulisan butir tes merupakan kegiatan menterjemahkan kisi-kisi dalam bentuk operasional. Untuk memperoleh hasil tes yang berkualitas diperlukan butir-butir yang baik dalam suatu perangkat tes sehingga dapat mengungkap kemampuan yang sebenarnya dan bukan kemampuan yang semu dari responden. Penulisan butir tes merupakan penciptaan kreatifitas yang menuntut kombinasi dari berbagai kemampuan yang dikembangkan melalui latihan, pengalaman, penguasaan bentuk tes, dan teknik penulisan butir tes (Susetyo, 2011 : 79 – 80).

Dalam penulisan butir soal, digunakan pedoman yang disampaikan oleh Susetyo, (2011 : 80) yaitu :

- 1) Penulisan setiap butir tes harus sejelas mungkin.
- 2) Menggunakan kata-kata yang mempunyai arti tepat.
- 3) Penggunaan kata-kata kompleks atau janggal perlu dihindari.
- 4) Berikan keterangan yang diperlukan secara lengkap untuk menjawab tes yang diberikan.
- 5) Penggunaan kata-kata yang tidak ada gunanya perlu dihindari.

- 6) Butir tes menggunakan rumusan kalimat yang setepat mungkin.
- 7) Taraf kesukaran butir tes perlu disesuaikan dengan kelompok dan tujuan pengetesan.
- 8) Perlu menghindari petunjuk ke arah jawaban benar yang tidak perlu.

Pada tahapan ini butir soal ditulis dari setiap dimensi berdasar kepada indikator-indikatornya. Setiap dimensi diwakili oleh minimal satu butir soal. Semakin luas materi yang terkandung dalam suatu dimensi maka semakin banyak pula butir soal yang mewakili dimensi tersebut.

#### **4. Penelaahan Butir Tes (Uji Validitas)**

Validitas dapat diartikan sejauhmana hasil pengukuran dapat diinterpretasikan sebagai cerminan sasaran ukur yang berupa kemampuan, karakteristik atau tingkah laku yang diukur melalui alat ukur yang tepat (Susetyo, 2011 : 89). Suatu tes dapat dikatakan mempunyai validitas yang tinggi apabila tes tersebut menjalankan fungsi ukurnya atau memberikan hasil ukur yang tepat dan akurat sesuai dengan maksud dikenakannya tes tersebut. Suatu tes yang menghasilkan data yang tidak relevan dengan tujuan

diadakannya pengukuran dikatakan sebagai tes yang memiliki validitas rendah (Azwar, 2013 : 173 – 174).

Pengujian validitas pada penelitian ini dilakukan sebelum alat

hli di bidang NKPI yang berkenan melakukan penelaahan butir tes pada penelitian ini.

ukur diujicobakan, yaitu melalui penilaian para ahli di bidang NKPI terhadap setiap butir tes. Berikut ini ialah para a

**Tabel 1. Penelaah Butir Tes**

No.	Nama	Profesi	Unit Kerja
1	Dr. Diyan Krisdiana, M. Si	Widyaiswara NKPI	PPPPTK Pertanian Cianjur
2	Muldan Martin, M.Si	Widyaiswara NKPI	PPPPTK Pertanian Cianjur
3	Vita Rumanti K, S.Pi, MT	Dosen TMPT	Institut Pertanian Bogor

Cara yang umum digunakan untuk mengetahui validitas isi adalah teknik kecocokan para ahli yang berkecimpung dalam bidang keilmuan tertentu (Susetyo, 2011 : 90). Setiap ahli melakukan penilaian dengan memberikan kecocokan atau ketidakcocokannya terhadap setiap butir tes berdasarkan indikatornya. Penilaian para ahli ini ditandai dengan pembubuhan tanda centang (√) pada kolom cocok atau kolom tidak cocok.

Terdapat beberapa teknik dalam melakukan perhitungan validitas isi, namun pada penelitian ini teknik yang

ian jumlah penilai yang merasa cocok akan dijumlahkan dan dipersentasekan terhadap jumlah penilai. Bila dirumuskan menjadi seperti berikut :

$$\text{Persentase} = \frac{f}{\sum f} \times 100\%$$

digunakan adalah dengan persentase butir yang cocok dengan indikator/tujuan. Perhitungan kecocokan terhadap validitas isi dilakukan dengan menghitung besarnya persentase pada pernyataan cocok, yaitu “persentase kecocokan suatu butir dengan tujuan/indikator” berdasarkan penilaian guru/dosen atau ahli (Noer, dalam Susetyo (2011 : 92)). Butir tes yang dianggap cocok oleh penilai diberi nilai 1 (satu), sedangkan butir tes yang dianggap tidak cocok oleh penilai diberi nilai 0 (nol). Kemudian

dimana :

$f$  = frekuensi cocok menurut penilai

$\Sigma f$  = jumlah penilai

Butir tes dinyatakan valid jika kecocokannya dengan indikator mencapai lebih besar dari 50% (Susetyo, 2011 : 92). Perhitungan kemudian dilakukan menggunakan software MS Excel. Dari hasil perhitungan, terdapat 5 butir soal yang persentasenya di bawah 50 % yaitu nomor soal 29, 30, 32, 49, dan 61. Sehingga dari 65 butir soal, hanya 60 butir soal yang dianggap valid dan dapat dilanjutkan ke tahap uji coba tes.

### 5. Uji Coba Tes 1

Pemilihan butir tes dapat dilakukan melalui ujicoba, dengan harapan akan memberikan umpan balik secara nyata dari responden secara langsung. Melalui hasil ujicoba akan diketahui kualitas butir tes yang telah dibuat apakah telah memenuhi persyaratan dan berkualitas, jika belum memenuhi kualitas dan persyaratan dilakukan revisi-revisi (Susetyo, 2011 : 81).

Mengingat perangkat tes ini ditujukan bagi calon-calon guru SMK Paket Keahlian NKPI dan syarat menjadi guru adalah lulusan S1/DIV, maka yang tepat untuk menjadi

responden uji coba pada penelitian ini adalah para mahasiswa tingkat akhir pada Program Studi Teknologi dan Manajemen Perikanan Tangkap (TMPT) (S1) dan guru-guru SMK Paket Keahlian NKPI.

Uji coba pertama dilakukan terhadap mahasiswa-mahasiswa Program Studi Teknologi dan Manajemen Perikanan Tangkap (TMPT) tingkat akhir. Uji coba dilakukan di kampus Institut Pertanian Bogor, Fakultas Perikanan dan Ilmu Kelautan. Pada uji coba pertama ini penulis hanya bisa mendapatkan 30 orang responden. Ketiga puluh orang mahasiswa ini diminta kesediaannya untuk menjawab 60 butir soal selama 30 menit.

### 6. Analisa Butir Tes 1

Analisa butir digunakan untuk menentukan kualitas butir pada suatu perangkat tes. Butir tes berkualitas rendah akan menurunkan kualitas ujian. Analisis butir pada penelitian ini akan dilakukan sesuai teori ujian klasik yang terdiri dari tingkat kesukaran, daya diskriminasi butir, dan efektivitas distraktor.

Tingkat kesukaran butir butir dengan benar dan banyaknya merupakan rasio antara penjawab penjawab butir. Formulasi tingkat kesukaran butir adalah :

$$p = n_i / N$$

dimana :

$n_i$  = banyaknya responden yang menjawab butir dengan benar

$N$  = banyaknya responden yang menjawab butir

Witherington (Susetyo, 2011 : 154) menjadi beberapa kategori seperti berikut.

**Tabel 2. Tingkat Kesukaran Menurut Witherington**

Rentang	Tingkat Kesukaran
$0,00 \leq P \leq 0,24$	Sukar
$0,25 \leq P \leq 0,74$	Sedang
$0,75 \leq P \leq 1,00$	Mudah

Daya diskriminasi butir adalah kemampuan butir dalam membedakan antara siswa yang memiliki kemampuan tinggi (Kelompok Unggul) dan siswa yang mempunyai kemampuan rendah (Kelompok Asor). Formulasi daya diskriminasi butir adalah :

$$d = (n_{iU} / N_U) - (n_{iA} / N_A)$$

dimana :

$n_{iU}$  = banyaknya responden dari kelompok unggul yang menjawab benar

$N_U$  = banyaknya responden dari kelompok unggul yang menjawab

$n_{iA}$  = banyaknya responden dari kelompok asor yang menjawab benar

$N_A$  = banyaknya responden dari kelompok asor yang menjawab

Menurut Nunnally dalam Susetyo (2011 : 161) koefisien korelasi di atas 0,20 sudah dianggap cukup baik dalam menetapkan daya beda pada suatu butir tes.

Efektifitas distraktor diperiksa untuk melihat apakah semua distraktor atau semua pilihan jawaban yang bukan kunci telah berfungsi sebagaimana mestinya, yaitu apakah distraktor-distraktor tersebut telah dipilih oleh banyak atau semua responden kelompok asor dan hanya sedikit atau tidak ada responden kelompok unggul yang memilihnya. Dalam memeriksa efektifitas distraktor umumnya lebih banyak menggunakan proporsi atau persentase jawaban bagi setiap alternatif sebagai ganti frekuensi jawaban dikarenakan proporsi dapat dibandingkan langsung antara satu butir dengan butir yang lain tanpa tergantung terhadap jumlah responden. Proporsi bagi setiap alternatif diperoleh dengan membagi frekuensi jawaban pada setiap pilihan dengan jumlah subjek dari masing-masing kelompok sehingga proporsi yang diperoleh merupakan proporsi dalam kelompok masing-masing (kelompok unggul dan kelompok asor).

Perhitungan tingkat kesukaran dan daya diskriminasi butir dilakukan secara manual dengan menggunakan *software* MS Excel. Sementara efektifitas distraktor dihitung dengan bantuan *software* ITEMAN.

Dari hasil analisa butir tes dapat diketahui bahwa dari 60 butir tes yang diujicobakan hanya 28 butir tes yang kemudian dianggap berkualitas. Dianggap berkualitas ditinjau dari daya diskriminasinya yang baik dan efektifitas distraktornya yang juga baik. Kartowagiran (2009 : 17) mengatakan bahwa sekalipun suatu butir tes terlalu sukar atau terlalu mudah, namun apabila (1) daya pembeda butir tes, dan (2) distribusi jawaban, memenuhi kriteria, maka butir tes tersebut masih dapat diterima sebagai butir tes yang baik. Namun untuk butir soal nomor 17 walaupun daya diskriminasinya tergolong tidak baik dan efektifitas distraktornya pun masuk pada kriteria jawaban terbalik, butir soal tersebut tetap dipertahankan karena mengingat butir tersebut satu-satunya butir yang mewakili dimensi Bahasa Inggris Maritim dan Perikanan. Sehingga ada 29 butir tes yang kemudian dirakit kembali menjadi sebuah perangkat tes dan siap digunakan untuk ujicoba kedua.

## 7. Uji Reliabilitas

Reliabilitas merupakan penerjemahan dari kata *reliability*. Suatu pengukuran yang mampu menghasilkan data yang memiliki tingkat reliabilitas tinggi disebut sebagai pengukuran yang reliabel

(*reliable*). Walaupun istilah reliabilitas mempunyai berbagai nama lain seperti konsistensi, keterandalan, keterpercayaan, kestabilan, keajegan, dan sebagainya, namun gagasan pokok yang terkandung dalam konsep reliabilitas adalah sejauhmana hasil suatu proses pengukuran dapat dipercaya. Hasil suatu pengukuran akan dapat dipercaya hanya apabila dalam beberapa kali pelaksanaan pengukuran terhadap kelompok subjek yang sama diperoleh hasil yang relatif sama, selama aspek yang diukur dalam diri subjek memang belum berubah (Azwar, 2012 : 7).

Koefisien reliabilitas alpha Cronbach digunakan untuk butir yang politomi, sehingga sering digunakan untuk tes berbentuk essay. Namun demikian koefisien reliabilitas jika menggunakan skor butir yang dikotomi, 1 dan 0, akan menjadi koefisien reliabilitas Kuder-Richardson. Koefisien reliabilitas Kuder Richardson tergolong koefisien korelasi batas bawah terutama pada rumus yang ke 20 (Susetyo, 2011 : 120). Oleh karena itu, dalam memperhitungkan nilai koefisien reliabilitas pada penelitian ini digunakan perhitungan reliabilitas Kuder Richardson (KR<sub>20</sub>).

Variansi butir yang diperlukan dalam perhitungan dapat disederhanakan menjadi :

$$\sigma_1^2 = p_1q_1 \quad \text{atau} \quad \Sigma\sigma_1^2 = \Sigma p_1q_1$$

apabila semua butir tes setara maka koefisien reliabilitas KR<sub>20</sub> adalah :

$$\rho_{KR20} = \frac{k}{k-1} \left[ 1 - \frac{\Sigma pq}{\sigma_A^2} \right] \quad \text{atau} \quad \rho_{KR20} = \frac{k}{k-1} \left[ \frac{\sigma_A^2 - \Sigma pq}{\sigma_A^2} \right]$$

dimana :

p = proporsi jawaban benar

q = proporsi jawaban salah

k = jumlah butir tes

N = jumlah responden

$\Sigma pq$  = jumlah perkalian jawaban benar dengan salah

$\rho_{KR20}$  = koefisien reliabilitas

$\sigma_A^2$  = varian skor tes

$$\sigma_A^2 = \frac{N \sum X^2 - (\sum X)^2}{N^2}$$

Suatu perangkat tes dinyatakan reliabel jika telah mencapai sekurang-kurangnya memperoleh koefisien korelasi sebesar 0,50. Di samping itu, besar kecilnya koefisien reliabilitas pada perangkat tes ditentukan juga oleh standar pada cabang keilmuan masing-masing (Susetyo, 2011 : 107).

Perhitungan reliabilitas dilakukan secara manual dengan menggunakan *software* MS Excel dan secara otomatis dengan bantuan *software* ITEMAN. Baik dari perhitungan dengan MS Excel maupun ITEMAN diperoleh hasil yang sama yaitu nilai koefisien reliabilitas sebesar 0,71. Nilai ini lebih besar dari 0,50 sehingga perangkat tes dapat dikategorikan reliabel.

### 8. Uji Coba Tes 2

Uji coba kedua dilakukan terhadap mahasiswa-mahasiswa Program Studi Teknologi dan Manajemen Perikanan Tangkap (TMPT) tingkat akhir, alumni-alumni Program Studi TMPT, dan guru-guru SMK Paket Keahlian NKPI. Uji coba dengan responden mahasiswa TMPT tingkat akhir dilakukan di kampus Institut Pertanian Bogor, Fakultas Perikanan dan Ilmu Kelautan dengan

jumlah responden 65 orang. Uji coba dengan responden alumni TMPT dilakukan melalui berkirim surat elektronik (*email*) dengan jumlah responden 10 orang. Uji coba dengan responden guru-guru SMK Paket Keahlian NKPI dilakukan di kampus Pusat Pengembangan dan Pemberdayaan Pendidik dan Tenaga Kependidikan (PPPPTK) Pertanian Cianjur dan SMK Negeri 1 Karang Tengah Cianjur dengan jumlah responden 25 orang. Sehingga jumlah seluruh responden adalah 100 orang. Keseratus orang responden ini diminta kesediaannya untuk menjawab 29 butir soal selama 15 menit.

### 9. Analisa Butir Tes 2

Terhadap hasil ujicoba tersebut akan dilakukan analisis butir untuk dijadikan dasar dalam memilah pertanyaan-pertanyaan mana saja yang berkualitas. Kualitas tes sangat ditentukan oleh kualitas butir-butirnya. Tes yang berkualitas tinggi dengan jumlah sedikit akan lebih berguna daripada tes yang berkualitas rendah dengan jumlah banyak. Sebab tes yang berkualitas rendah akan memberikan hasil pengukuran yang menyesatkan.

Seperti halnya pada analisa butir tes pertama, analisa butir tes kedua ini pun akan dilakukan sesuai teori ujian klasik yang terdiri dari tingkat kesukaran, daya diskriminasi butir, dan efektivitas distraktor.

Perhitungan tingkat kesukaran dan daya diskriminasi butir dilakukan secara manual dengan menggunakan *software* MS Excel. Sementara efektivitas distraktor dihitung dengan bantuan *software* ITEMAN.

Dari hasil analisa butir tes tersebut dapat diketahui bahwa dari 29 butir tes diujicobakan hanya 19 butir tes yang kemudian dianggap berkualitas. Dianggap berkualitas ditinjau dari daya diskriminasinya yang baik dan efektivitas distraktornya yang juga baik. Kartowagiran (2009 : 17) mengatakan bahwa sekalipun suatu butir tes terlalu sukar atau terlalu mudah, namun apabila (1) daya pembeda butir tes, dan (2) distribusi jawaban, memenuhi kriteria, maka butir tes tersebut masih dapat diterima sebagai butir tes yang baik. Bila pada analisa butir pertama, butir yang memiliki daya diskriminasinya tergolong tidak baik dan efektifitas distraktornya pun masuk pada kriteria jawaban terbalik masih dipertimbangkan untuk tetap dimasukkan mengingat butir tersebut

satu-satunya butir yang mewakili dimensinya, namun pada analisa butir kedua ini hal tersebut tidak dilakukan kembali demi mempertahankan kualitas perangkat tes. Sehingga hanya 19 butir tes yang berkualitas baik yang dirakit kembali menjadi sebuah perangkat tes bentuk akhir.

#### **10. Perakitan Bentuk Akhir Tes**

Setelah perangkat tes melalui proses uji validitas, ujicoba, analisa butir, uji reliabilitas, kemudian diujicobakan kembali dan dilakukan analisa butir lagi, maka diperoleh tes bentuk akhir dengan jumlah butir tersisa sebanyak 19 butir tes. Bila dibandingkan dengan jumlah awalnya sebanyak 65 butir tes, berarti hanya 29% saja yang lolos seleksi hingga akhir. Namun demikian, 19 butir tes yang tersisa tersebut telah lulus uji hingga dapat dikatakan sebagai soal yang berkualitas baik, dalam hal ini soal yang valid dan reliabel.

#### **C. SIMPULAN DAN REKOMENDASI**

Dari pembahasan pada bab sebelumnya dapat disimpulkan bahwa :

Pengembangan perangkat tes melalui prosedur yang baku dapat menghasilkan perangkat tes yang valid dan reliabel. Prosedur yang dilalui dalam penelitian ini yaitu : identifikasi tujuan pengukuran,

pengembangan spesifikasi tes, penulisan butir soal, penelaahan butir tes (uji validitas), uji coba tes 1, analisa butir tes 1, uji reliabilitas, uji coba tes 2, analisa butir tes 2, dan perakitan bentuk akhir tes.

Pada awal pembuatan spesifikasi tes dan penulisan butir, jumlah butir yang dikembangkan berjumlah 65 butir tes. Seiring proses pengembangan tes dilakukan, jumlah butir tes mengalami degradasi (penurunan). Setelah melalui tahap telaah butir (uji validitas) yang dilakukan oleh 3 orang ahli di bidang NKPI, butir tes mengalami pengurangan 5 butir tes sehingga tersisa 60 butir tes untuk dirakit menjadi perangkat tes uji coba 1. Setelah uji coba 1 dilakukan dilanjutkan dengan analisis butir 1. Pada tahapan ini butir soal pun mengalami pengurangan hingga tersisa 29 butir tes untuk dirakit menjadi perangkat tes uji coba 2. Uji coba 2 dilakukan dan diteruskan dengan analisa butir 2, kembali lagi butir soal mengalami pengurangan hingga tersisa 19 butir tes untuk dirakit menjadi perangkat tes bentuk akhir.

Perhitungan reliabilitas dilakukan secara manual dengan menggunakan *software* MS Excel dan secara otomatis dengan bantuan *software* ITEMAN. Estimasi reliabilitas yang dilakukan secara manual dalam penelitian ini menggunakan pendekatan konsistensi internal (*internal consistency*) yaitu dengan perhitungan

reliabilitas Kuder Richardson ( $KR_{20}$ ). Baik dari perhitungan dengan MS Excel maupun ITEMAN diperoleh hasil yang sama. Nilai koefisien reliabilitas diperoleh nilai sebesar 0,71. Nilai ini lebih besar dari 0,50 sehingga perangkat tes tergolong reliabel.

#### D. DAFTAR PUSTAKA

- Ali, M. (2011). *Memahami Riset Prilaku dan Sosial*. Bandung : Pustaka Cendekia Utama.
- Anderson, L. dan Krathwohl, D.R. (2010). *Kerangka Landasan Untuk Pembelajaran, Pengajaran, dan Asesmen (Revisi Taksonomi Pendidikan Bloom)*. Yogyakarta : Pustaka Pelajar.
- Anonim. (2008). *Peraturan Pemerintah Nomor 74 Tahun 2008 Tentang Guru*. Diakses dari website <http://www.setneg.go.id>
- Azwar, S. (2013). *Tes Prestasi : Fungsi dan Pengembangan Pengukuran Prestasi Belajar* (Edisi Kedua). Yogyakarta : Pustaka Pelajar.
- Azwar, S. (2012). *Relibilitas dan Validitas* (Edisi Keempat). Yogyakarta : Pustaka Pelajar.
- Kartowagiran, B. (2009). *Pengantar Teori Tes Klasik*. Makalah. Yogyakarta : Pasca Sarjana UNY.

Susetyo, B. (2011). *Menyusun Tes Hasil Belajar*. Bandung : CV Cakra.

Universitas Pendidikan Indonesia. (2013).  
*Pedoman Penulisan Karya Ilmiah*.  
Bandung : UPI PRESS.

Zainul, A. dan Nasoetion, N. (1993).  
*Penilaian Hasil Belajar*. Jakarta :  
PAU-PPAI Universitas Terbuka.