**PENGEMBANGAN INSTRUMEN PENILAIAN HASIL BELAJAR PKn**

**PESERTA DIDIK KELAS XI IPA DI SMAN 1 DONRI DONRI**

**KABUPATEN SOPPENG**

**Oleh: Nuraeni**

Universitas Negeri Makassar Jl. Bonto Langkasa, Makassar 90222, Telp (0411) 830366, Fax (0411) 855288, E-mail: [pasca@unm.ac.id](mailto:pasca@unm.ac.id).

Website: http://www.pps.unm.ac.id

**Abstrak**

Penelitian ini bertujuan untuk mendeskripsikan hasil pengembangan instrumen hasil belajar PKn pada peserta didik Kelas XI IPA SMA Negeri 1 Donri Donri Kabupaten Soppeng ditinjau dari validitas isi (*content validity*), tingkat kesukaran, daya beda butir, efektivitas pengecoh, validitas butir, dan reliabilitas. Penelitian ini jenis penelitian dan pengembangan (*Research & Development*) dengan subjek penelitian ini ialah peserta didik Kelas XI IPA SMA Negeri 1 Donri Donri Kabupaten Soppeng. Pengumpulan data dilakukan dengan tes. Teknik analisis data yang digunakan adalah kuantitatif deskriptif. Hasil penelitian menunjukkan bahwa instrumen penilaian hasil belajar PKn di SMA Negeri 1 Donri-Donri yang dikembangakan sebanyak 61 butir telah memenuhi syarat. Instrumen penilaian hasil belajar PKn di SMA Negeri 1 Donri-Donri yang baku sebanyak 61 butir memiliki sebaran tingkat kesukaran butir yaitu sebanyak 2 butir berada pada kategori sulit, 40 butir berada pada kategori sedang, dan 19 butir berada pada kategori mudah dengan rata-rata tingkat kesukaran sebesar 0.6. Instrumen penilaian hasil belajar PKn di SMA Negeri 1 Donri-Donri yang baku sebanyak 61 butir telah memiliki daya beda butir yang baik dengan rata-rata indeks daya beda butir sebesar 0.68. Instrumen penilaian hasil belajar PKn di SMA Negeri 1 Donri-Donri yang baku sebanyak 61 butir telah memiliki pengecoh yang berfungsi. Instrumen penilaian hasil belajar PKn di SMA Negeri 1 Donri-Donri yang baku sebanyak 61 butir telah valid secara secara empirik berdasarkan uji coba lapangan dengan rata-rata indeks validitas sebesar 0.52. Instrumen penilaian hasil belajar PKn di SMA Negeri 1 Donri-Donri yang baku sebanyak 61 butir telah reliabel dengan indeks reliabilitas sebesar 0.95.

***Abstract***

*This study aims at describing the result of PKn (Civics) instrument development of class XI IPA students at SMAN 1 Donri Donri in Soppeng disrict based on content validity, level of difficulty, item discrimination, deceiver effectiveness, item validity, and reliability. The study is a research and development. The subject were class XI IPA students at SMAN 1 Donri Donri in Soppeng district. Data was collected bay employing test. Data was analyzed using descriptive quantitative analysis. The result of the study reveals that the PKn assessment instrument at SMAN 1 Donri Donri which is developed are 61 items with items difficulties, namely 2 items are in difficult category, 40 items are in medium category, and 19 items are in easy category. The standard PKn assessment instrument at SMAN 1 Donri Donri is 61 items which have good discrimation. The standard PKn assessment instrument at SMAN 1 Donri Donri is 61 items which have funtional deceiver. The standard PKn assessment instrument at SMAN 1 Donri Donri is 61 items which are valid empirically based on the field test. The standard PKn assessment instrument at SMAN 1 Donri Donri is 61 items which are reliable.*

**Pendahuluan**

Pelaksanaan penilaian pada mata pelajaran PKn selama ini masih belum maksimal. Salah satu hal yang menjadi sorotan peneliti bahwa guru-guru dilapangan tampaknya belum mengerti terkait prosedur pengembangan instrumen penilaian hasil belajar yang benar. Fenomena yang terjadi di lapangan menunjukkan bahwa dalam mengevaluasi hasil belajar peserta didik, guru cenderung hanya mengembangkan instrumen penilaian yang seadanya. Bahkan sering ditemukan instrumen penilaian yang dikembangkan oleh guru hanya mengcopy dari buku paket saja tanpa adanya usaha mengembangkan sendiri melalui prosedur yang tepat.

Langkah-langkah pengembangan instrumen yang dilakukan guru selama ini hanya sebatas penyusunan kisi-kisi, pebuatan soal dan langsung diberikan kepada peserta didik. Tidak ada usaha guru selama ini untuk mempertimbangkan kualitas instrumen yang dikembangkan sebelum diserahkan kepada peserta didik. Atas dasar hal tersebut penulis menyadari bahwa hal tersebut adalah kekeliruan yang dialami selama ini. Hal lain yang dirasakan sebagai kendala selama ini adalah kurangnya pemahaman guru terkait teknik mengidentifikasi kualitas instrumen hasil belajar yang benar. Bahkan berdasarkan wawancara terbatas dengan guru di lapangan menunjukkan bahwa guru belum memahami karakteristik suatu instrumen sehingga dapat dikategorikan layak untuk digunakan dalam penyelenggaraan evaluasi pembelajaran PKn. Oleh karena itu, tidak heran jika instrumen penilaian hasil belajar PKn yang dikembangkan selama ini tidak jelas kualitasnya baik secara kualitatif maupun kualitas secara kuantitatif.

Permasalahan yang telah dikemukakan sebelumnya, maka peneliti menyimpulkan bahwa perlunya tindak lanjut atas permasalahan yang dialami oleh guru-guru PKn selama ini yaitu kurang maksimalnya penyelenggaraan evaluasi pembelajaran PKn di SMA selama ini. Tindak lanjut yang dimaksudkan dalam hal ini ialah upaya perbaikan pelaksanaan evaluasi pembelajaran selama ini. Salah satu wujud dari upaya perbaikan tersebut ialah pengembangan instrumen penilaian hasil belajar PKn yang valid dan reliabel berdasarkan prosedur pengembangan tes hasil belajar, sehingga dengan adanya pengembangan instrumen penilaian hasil belajar tersebut, diharapkan lahirnya instrumen yang berkualitas yang dapat menunjang penyelenggaraan evaluasi pembelajaran PKn yang berkualitas pula.

Berdasarkan hal-hal yang telah diuraikan tersebut maka peneliti bermaksud melakukan kajian teoretis dan empiris dalam memecahkan permasalahan yang dialami selama ini khususnya pada pelaksanaan evaluasi pembelajaran PKn. Upaya tersebut diwujudkan dalam suatu penelitian untuk mengembangkan instrumen penilaian hasil belajar PKn yang berkualitas. Oleh karena itu, pada penelitian ini peneliti mengangkat judul penelitian yaitu “Pengembangan Instrumen Penilaian Hasil Belajar PKn Peserta Didik Kelas XI IPA SMAN 1 Donri Donri Kabupaten Soppeng”.

Berdasarkan uraian latar belakang masalah yang telah dikemukakan sebelumnya, maka pertanyaan penelitian yang akan dijawab pada pelaksanaan penelitian ini sebagai berikut.

1. Bagaimanakah hasil pengembangan instrumen hasil belajar PKn berdasarkan hasil uji coba pada peserta didik Kelas XI IPA SMA Negeri 1 Donri Donri Kabupaten Soppeng ditinjau dari validitas isi (*content validity*)?
2. Bagaimanakah hasil pengembangan instrumen hasil belajar PKn berdasarkan hasil uji coba pada peserta didik Kelas XI IPA SMA Negeri 1 Donri Donri Kabupaten Soppeng ditinjau dari tingkat kesukaran butir?
3. Bagaimanakah hasil pengembangan instrumen hasil belajar PKn berdasarkan hasil uji coba pada peserta didik Kelas XI IPA SMA Negeri 1 Donri Donri Kabupaten Soppeng ditinjau dari daya beda butir?
4. Bagaimanakah hasil pengembangan instrumen hasil belajar PKn berdasarkan hasil uji coba pada peserta didik Kelas XI IPA SMA Negeri 1 Donri Donri Kabupaten Soppeng ditinjau dari efektivitas distraktor?
5. Bagaimanakah hasil pengembangan instrumen hasil belajar PKn berdasarkan hasil uji coba pada peserta didik Kelas XI IPA SMA Negeri 1 Donri Donri Kabupaten Soppeng ditinjau dari validitas butir?
6. Bagaimanakah hasil pengembangan instrumen hasil belajar PKn berdasarkan hasil uji coba pada peserta didik Kelas XI IPA SMA Negeri 1 Donri Donri Kabupaten Soppeng ditinjau dari reliabilitas?

Berdasarkan pertanyaan penelitian yang telah dikemukakan tersebut, maka tujuan pelaksanaan penelitian ini yaitu untuk mendeskripsikan hal-hal sebagai berikut.

1. Hasil pengembangan instrumen hasil belajar PKn berdasarkan hasil uji coba pada peserta didik Kelas XI IPA SMA Negeri 1 Donri Donri Kabupaten Soppeng ditinjau dari validitas isi (*content validity*).
2. Hasil pengembangan instrumen hasil belajar PKn berdasarkan hasil uji coba pada peserta didik Kelas XI IPA SMA Negeri 1 Donri Donri Kabupaten Soppeng ditinjau dari tingkat kesukaran butir.
3. Hasil pengembangan instrumen hasil belajar PKn berdasarkan hasil uji coba pada peserta didik Kelas XI IPA SMA Negeri 1 Donri Donri Kabupaten Soppeng ditinjau dari daya beda butir.
4. Hasil pengembangan instrumen hasil belajar PKn berdasarkan hasil uji coba pada peserta didik Kelas XI IPA SMA Negeri 1 Donri Donri Kabupaten Soppeng ditinjau dari efektivitas distraktor.
5. Hasil pengembangan instrumen hasil belajar PKn berdasarkan hasil uji coba pada peserta didik Kelas XI IPA SMA Negeri 1 Donri Donri Kabupaten Soppeng ditinjau dari validitas butir.
6. Hasil pengembangan instrumen hasil belajar PKn berdasarkan hasil uji coba pada peserta didik Kelas XI IPA SMA Negeri 1 Donri Donri Kabupaten Soppeng ditinjau dari reliabilitas.

.

Manfaat yang diharapkan dari pelaksanaan penelitian ini adalah:

* 1. Manfaat Teoretis

Sebagai sarana untuk menerapkan teori pengembangan instrumen hasil belajar pada Mata Pelajaran PKn di Kelas XI IPA SMA Negeri 1 Donri Donri Kabupaten Soppeng. Selain itu, penelitian ini juga diharapkan bermanfaat sebagai bahan informasi mengenai karakteristik instrumen hasil belajar pada Mata Pelajaran PKn serta sebagai bank soal yang dapat digunakan dimasa yang akan datang.

* 1. Manfaat Praktis

Bagi MGMP PKn Kabupaten Soppeng, sebagai bahan masukan untuk pengembangan instrumen hasil belajar pada Mata Pelajaran PKn di Kelas XI IPA SMA Negeri 1 Donri Donri Kabupaten Soppeng. Bagi Dinas Pendidikan Kabupaten Soppeng, sebagai dasar dalam menentukan kebijakan untuk pelaksanaan evaluasi pembelajaran di SMA.

**Metode Penelitian**

Penelitian ini jenis penelitian dan pengembangan (*Research & Development*) dengan subjek penelitian ini ialah peserta didik Kelas XI IPA SMA Negeri 1 Donri Donri Kabupaten Soppeng. Pengumpulan data dilakukan dengan tes. Teknik analisis data yang digunakan adalah kuantitatif deskriptif. Untuk memberikan interpretasi terhadap hasil analisis yang diperoleh, maka berikut ini disajikan kriteria kualitas instrumen yang dijadikan pedoman dalam pengembangan instrumen penilaian hasil belajar PKn:

1. Validitas Isi

Menurut Lawshe dan Martuza (dalam Ruslan: 2009) membahas metode statistika untuk menentukan validitas isi dan realibilitas menyeluruh dari suatu tes melalui penilaian pakar. Relevansi kedua pakar secara menyeluruh merupakan validitas isi Gregory, yaitu koefisien validitas isi. Untuk memutuskan apakah instrumen telah memiliki derajat validitas yang memadai, maka digunakan model kesepakatan tersebut dengan kriteria hasil penilaian dari kedua validator minimal memiliki “relevansi kuat”. Jika hasil dari koefisien validitas isi ini tinggi (V > 75%), maka hasil pengukuran yang dilakukan adalah sahih.

1. Daya beda

Menurut Crocker & Algina (Mansyur, Rasyid dan Suratno, 2009) bahwa kriteria besarnya koefisien daya beda yaitu:

0,4 ≤ D ≤ 1,0 = Baik

0,3 ≤ D ≤ 0,39 = Dapat diterima tanpa revisi

0,2 ≤ D ≤ 0,29 = Dapat diterima dengan revisi

D < 0,2 = Buruk (tidak berfungsi).

1. Tingkat kesukaran

Menurut Mansyur, Rasyid dan Suratno (2009), kriteria yang digunakan untuk menentukan jenis tingkat kesukaran butir soal adalah sebagai berikut:

P ≤ 0,30 = butir soal sukar

0,3 < P ≤ 0,70 = butir soal sedang

P > 0,70 = butir soal mudah

1. Efektivitas Pengecoh

Menurut Mardapi (2008), bahwa penyebaran pilihan jawaban dijadikan dasar dalam penelaahan soal. Suatu pilihan jawaban atau pengecoh dapat dikatakan berfungsi apabila pengecoh paling tidak dipilih oleh 5 % peserta tes.

1. Validitas Butir

Suatu butir dikatakan valid apabila memiliki sumbangan terhadap skor totalnya. Jika koofisien korelasi butir terhadap skor totalnya tersebut lebih atau sama dengan 0.3, maka butir tersebut dikatakan valid secara empirik berdasarkan kriteria internalnya.

1. Reliabilitas

Menurut Mansyur, Rasyid dan Suratno (2009) bahwa = korelasi antara observed-scores (skor amatan) dari suatu tes. Menurut Linn (1989) dalam Mansyur, Rasyid dan Suratno (2009) nilai koefisien reliabilitas termasuk tinggi, apabila batas minimal suatu tes memiliki keajegan (konsisten dan stabil) adalah 0,70 sedangkan suatu tes dinyatakan tidak reliabel apabila batas suatu tes < 0,70.

**Hasil Penelitian dan Pembahasan**

* + 1. **Hasil Pengembangan Instrumen berdasarkan Validitas Isi**

Hasil penelitian yang diperoleh dari pengembangan instrumen hasil belajar PKn di SMA Negeri 1 Donri-Donri berdasarkan aspek kualitas validitas isi dari penilaian pada instrumen tes tersebut pada tingkat relevansi antara standar kompetensi, kompetensi dasar, dan indikator, terhadap butir-butir soal. Berdasarkan hasil penilaian tingkat relevansi oleh pakar yang telah ditunjuk berdasarkan pertimbangabn keahlian pada bidang PKn tersebut diperoleh hasil penelitian berupa tingkat relevansi oleh pakar terhadap instrumen yang dikembangkan sebesar 0.95.

Jika temuan empirik tersebut dikaitkan dengan konsep teoretis sebagai mana yang dikemukakan oleh Gregory (Ruslan 2007) jika indeks V > 0.75 maka instrumen dinyatakan valid secara isi. Berdasarkan indeks kesepakatan pakar (0.95) tersebut menunjukkan bahwa instrumen hasil belajar PKn yang telah dirancang memenuhi kriteria dalam validitas isi. Fitzpatric (1983: 5) menjelaskan bahwa validitas isi menunjukkan ketepatan sampling butir dapat mengukur konten pengujiannya. Lebih lanjut Sukardi (2011: 31), validitas suatu instrumen penilaian, tidak lain adalah derajat yang menunjukkan di mana suatu tes mengukur apa yang hendak diukur yang memilki beberapa makna penting. Salah satu hal penting dalam instrumen penilaian adalah sejauh mana instrumen penilaian tersebut mampu mengukur kontennya atau dalam hal ini adalah standar kompetensi, kompetensi dasar dan inikator soal dari suatu materi yang ingin diukur. Untuk itu berdasarkan temuan empirik dan konsep teoretis berdasarkan validitas isi instrumen, maka dapat disimpulkan bahwa instrumen penilaian hasil belajar PKn SMA Negeri 1 Donri-Donri sudah mampu mengukur aspek kontennya dengan baik, yaitu standar kompetensi, kompetensi dasar, dan indikator.

* + 1. **Hasil Pengembangan Instrumen berdasarkan Tingkat Kesukaran**

Hasil penelitian dari pengembangan instrumen penilaian hasil belajar PKn di SMA Negeri 1 Donri-Donri berdasarkan tingkat kesukaran butir diperoleh temuan empirik bahwa dari 62 butir instrumen hasil belajar PKn hasil revisi uji coba skala kecil, sebanyak 2 butir soal yang berada pada kategori sulit, sebanyak 40 butir soal yang berada dalam kategori sedang, serta sebanyak 20 butir soal yang berada pada kategori mudah. Namun hal lain yang perlu dijelaskan bahwa salah satu butir pada instrumen tersebut yang memiliki kualitas buruk untuk daya beda dan validitas. Butir tersebut adalah butir 50 yang digugurkan, oleh karena itu instrumen yang baku total sebanyak 61 butir. Dari 61 butir instrumen yang telah baku tersebut sebanyak 2 butir soal yang berada pada kategori sulit, sebanyak 40 butir soal yang berada dalam kategori sedang, serta sebanyak 19 butir soal yang berada pada kategori mudah.

Jika temuan empirik tersebut dikaitkan dengan tinjauan teoretis sebagaimana yang dikemukakan oleh Azwar (2010: 134) menyatakan bahwa indeks kesukaran butir adalah rasio penjawab butir dengan benar dan banyaknya penjawab butir. Tingkat kesukaran butir soal dapat digunakan sebagai pertimbangan untuk memilih butir soal dalam suatu bank soal sesuai dengan kebutuhan evaluator yang ingin melakukan penilaian. Untuk itu pada penelitian ini telah membakukan 61 butir soal PKn untuk Kelas XI IPA yang siap digunakan oleh guru PKn untuk keperluan penilaian.

* + 1. **Hasil Pengembangan Instrumen berdasarkan Daya Beda Butir**

Pada aspek daya beda butir, hasil penelitian ini meneunjukkan bahwa dari 62 butir instrumen hasil belajar PKn hasil revisi uji coba skala kecil, sebanyak 61 butir soal yang berada pada kategori daya beda baik, sebanyak 0 butir soal yang berada dalam kategori daya beda cukup baik, sebanyak 0 butir soal yang berada pada kategori daya beda kurang baik, dan sebanyak 1 butir soal yang berada pada kategori daya beda buruk. Sehingga hasil tersebut menyimpulkan bahwa instrumen penilaian hasil belajar PKn yang baku dalam penelitian ini sebanyak 61 butir jika ditinjau dari aspek daya beda butir.

Ditinjau dari aspek teoretis yang telah dibangun pada kajian teoretis sebagaimana dikemukakan oleh Naga (1992:67), bahwa makin tinggi daya pembedaan butir, maka makin besar perbedaan skor yang dihasilkan oleh kelompok tinggi dan kelompok rendah. Fungsi dari daya pembeda tersebut adalah mendeteksi perbedaan individual yang sekecil-kecilnya di antara para peserta tes. Crocker & Algina dalam Surapranata (2004:24) mendefinisikan indeks daya pembeda sebagai selisih antara proporsi jawaban benar pada kelompok atas dengan-proporsi jawaban benar pada kelompok bawah. Boopathiraj & Chellamani (2013:190) menjelaskan bahwa daya beda menunjukkan kekuatan tes dan butir dalam membedakan kemampuan peserta tes. Peserta yang berkemampuan tinggi seharusnya menjawab benar dan peserta yang berkemampuan rendah idealnya akan salah.

Jika temuan empirik pada pelaksanaan penelitian ini dikaitkan dengan tinjauan teoretis terkait dengan konsep daya beda butir, maka dapat disimpulkan bahwa 61 butir instrumen penilaian hasil belajar PKn yang dikemukakan pada sebelumnya sudah dapat membedakan antara kelompok bawah dengan kelompok atas. Dengan kata lain 61 butir instrumen penilaian hasil belajar PKn tersebut sudah mampu membedakan antara peserta didik yang menguasai materi PKn, dengan peserta didik yang kurang menguasai materi PKn.

* + 1. **Hasil Pengembangan Instrumen berdasarkan Efektifitas Pengecoh**

Hasil pengembangan instrumen penilaian hasil belajar PKn di SMA Negeri 1 Donri-Donri dalam penelitian ini diperoleh temuan empirik bahwa dari 61 butir soal yang telah baku pada penelitian ini memiliki distraktor atau pengecoh yang berfungsi. Dengan kata lain semua butir instrumen yang telah baku memiliki pengecoh yang berfungsi. Hal tersebut diperoleh setelah melihat fungsi pengecoh dari instrumen penilaian hasil belajar PKn yang telah diuji cobakan ditemukan bahwa semua butir soal mengandung pengecoh yang berfungsi.

Secara teori, Menurut Mardapi (2008), bahwa penyebaran pilihan jawaban dijadikan dasar dalam penelaahan soal. Hal ini dimaksudkan untuk mengetahui berfungsi tidaknya jawaban yang tersedia. Suatu pilihan jawaban atau pengecoh dapat dikatakan berfungsi apabila pengecoh paling tidak dipilih oleh 5 % peserta tes/ peserta didik. Hal serupa dikemukakan Ali,S. & Khaeruddin (2012: 96) bahwa “Distraktor sudah berfungsi dengan baik jika sudah dipilih paling sedikit 5% dari peserta tes (p > 5%). Sudijono (2012: 411) mengemukakan bahwa, “Distraktor dinyatakan telah dapat menjalankan fungsinya dengan baik apabilah distraktor tersebut sekurang-kurangnya sudah dipilih oleh 5% dari seluruh peserta tes”.

Jika temuan empirik penelitian ini dikaitkan dengan tinjauan teoretis, maka semua distraktor atau pengecoh pada 61 butir instrumen penilaian hasil belajar PKn yang telah baku sudah teruji dan dapat menjalan fungsinya sebagai pengecoh. Fungsi pengecoh dalam hal ini adalah pengecoh bagi peserta didik yang kurang memahami materi PKn yang sedang diujikan.

* + 1. **Hasil Pengembangan Instrumen berdasarkan Validitas Butir**

Hasil pengembangan instrumen penilaian hasil belajar Pkn yang diperoleh dari pelaksanaan penelitian ini ditemukan bahwa sebanyak 61 butir yang telah valid secara empirik. Selain itu terdapat 1 butir soal yang gugur karena tidak valid. Butir soal yang gugur tersebut ialah butir 50. Dari hasil analisis tersebut dapat diputuskan bahwa butir yang baku dalam aspek validitas butir pada pengembangan instrumen hasil belajar PKn ini sebanyak 61 butir soal.

Secara teoretis sebagaimana yang dikemukakan oleh Mansyur, Rasyid dan Suratno (2009: 252) bahwa konsep validitas butir dengan konsistensi internal dapat dilakukan dengan mengkorelasikan skor butir dengan skor totalnya. Hal serupa dikemukakan oleh Widoyoko (2012: 147) bahwa suatu butir dikatakan valid apabila memiliki sumbangan yang besar terhadap skor totalnya. Untuk itu butir yang valid memiliki kemampuan untuk mengukur apa yang hendak diukur sebagaimana yang dicerminkan pada skor total peserta.

Jika temuan empirik tersebut dikaitkan dengan konsep teoretis terkait validitas butir tersebut, maka dapat disimpulkan bahwa 61 butir instrumen penilaian hasil belajar PKn yang telah baku dalam proses pengembangan ini telah memiliki kemampuan untuk mengukur apa yang hendak diukur sebagaimana yang dicerminkan pada skor total peserta. Dengan demikian hasil tersebut menunjukkan bahwa instrumen penilaian hasil belajar PKn ini sudah dapat digunakan untuk mengukur kompetensi peserta didik dalam mata pelajaran PKn.

* + 1. **Hasil Pengembangan Instrumen berdasarkan Reliabilitas**

Hasil penelitian secara empirik pada proses pengembangan instrumen hasil belajar PKn di SMA Negeri 1 Donri-Donri ditinjau dari aspek reliabilitas instrumen dikemukakan bahwa berdasarkan hasil penelitian pada proses pengembangan instrumen, maka diperoleh koofisien reliabilitas sebesar 0.95. Secara teoretis reliabilitas merupakan terjemahan dari kata *realibility* yang mempunyai asal kata *rely* dan *ability* (Azwar, 2010). Pengukuran yang memiliki realibilitas tinggi dikatakan sebagai pengukuran yang reliabel. Reliabilitas sendiri mempunyai nama lain seperti keterpercayaan, keterandalan, keajegan, kestabilan, konsistensi, dan sebagainya. Namun demikian, ide pokok yang terkandung dalam konsep realibilitas adalah sejauhmana hasil suatu pengukuran dapat dipercaya. Ziegler & Detje (2012: 3) menjelaskan bahwa reliabilitas menggambarkan konsistensi keseluruhan pengukuran meskipun diberikan beberapa kali. Hasil pengukuran dapat dipercaya hanya apabila dalam beberapa kali pelaksanaan pengukuran terhadap kelompok subyek yang sama diperoleh hasil yang relatif sama, selama aspek yang diukur dalam diri subyek memang belum berubah (Azwar, 2010). Dalam pengertian ini, relatif sama berarti tetap adanya toleransi terhadap perbedaan-perbedaan kecil diantara hasil beberapa kali pengukuran. Bila perbedaan itu sangat besar dari waktu ke waktu maka hasil pengukuran tidak dapat dipercaya dan dikatakan sebagai tidak reliabel.

Menurut Mansyur, Rasyid dan Suratno (2009) bahwa = korelasi antara observed-scores (skor amatan) dari suatu tes. Menurut Linn (1989) dalam Mansyur, Rasyid dan Suratno (2009) nilai koefisien reliabilitas termasuk tinggi, apabila batas minimal suatu tes memiliki keajegan (konsisten dan stabil) adalah 0,70, sedangkan suatu tes dinyatakan tidak reliabel apabila batas suatu tes < 0,70. Untuk itu jika temuan empirik pada penelitian ini dikaitkan dengan konsep teoretis berdasarkan aspek reliabilitas, maka dapat disimpulkan bahwa instrumen penilaian hasil belajar PKn di SMA Negeri 1 Donri-Donri Kabupaten Soppeng telah andal atau reliabel dan jika digunakan untuk keperluan pengukuran hasil belajar, maka hasil pengukuran yang diperoleh dapat dipercaya.

**Kesimpulan**

Berdasarkan hasil penelitian dan pembahasan, maka diperoleh kesimpulan penelitian sebagai berikut.

1. Instrumen penilaian hasil belajar PKn di SMA Negeri 1 Donri-Donri yang dikembangkan sebanyak 61 butir telah valid secara isi.
2. Instrumen penilaian hasil belajar PKn di SMA Negeri 1 Donri-Donri yang dikembangkan sebanyak 61 butir memiliki sebaran tingkat kesukaran butir yaitu sebanyak 2 butir berada pada kategori sulit, 40 butir berada pada kategori sedang, dan 19 butir berada pada kategori mudah. Rata-rata tingkat kesukaran sebesar 0.6.
3. Instrumen penilaian hasil belajar PKn di SMA Negeri 1 Donri-Donri yang dikembangkan sebanyak 61 butir telah memiliki daya beda butir yang baik. Rata-rata indeks daya beda sebesar 0.68.
4. Instrumen penilaian hasil belajar PKn di SMA Negeri 1 Donri-Donri yang dikembangkan sebanyak 61 butir telah memiliki pengecoh yang berfungsi.
5. Instrumen penilaian hasil belajar PKn di SMA Negeri 1 Donri-Donri yang dikembangkan sebanyak 61 butir telah valid secara secara empirik berdasarkan uji coba lapangan dengan rata-rata indeks validitas sebesar 0.52.
6. Instrumen penilaian hasil belajar PKn di SMA Negeri 1 Donri-Donri yang dikembangkan sebanyak 61 butir telah reliabel dengan indeks reliabilitas 0.95.

**Saran**

Berdasarkan kesimpulan yang diperoleh dalam penelitian ini, maka disarankan hal-hal berikut ini.

* 1. Kepada guru-guru PKn, disarankan untuk menggunakan instrumen hasil belajar PKn yang telah dikembangkan, karena instrumen tersebut telah mmiliki kualitas yang baik berdasarkan proses pengembangan instrumen.
  2. Kepada kepala sekolah agar senantiasa memantau guru-guru dalam melakukan penilaian hasil belajar, khususnya terkait dengan perkembangan instrumen penilaian hasil belajar yang benar agar hasil pengukuran dan penilaian yang diperoleh dapat lebih objektif dan dapat dipertanggungjawabkan.
  3. Kepada Kepala Dinas Pendidikan Kabupaten Soppeng agar memfasilitasi guru-guru dalam pelatihan pengembangan instrumen dengan mendatangkan pakar-pakar dibidang tersebut, agar kualitas penilaian di Kabupaten Soppeng bisa lebih baik di masa yang akan datang.

**Daftar Pustaka**

Ali, S. & Khaeruddin. 2012. *Evaluasi Pembelajaran.* Makassar: Badan Penerbit UNM.

Aryani, Ine Kusuma dan Susantim, Markum. 2010. *Pendidikan Kewarganegaraan Berbasis Nilai*. Bogor: Ghalia Indonesia

Azwar, S. 2010. *Tes Prestasi (Fungsi dan Pengembangan Pengukuran Prestasi Belajar)*. Yogyakarta: Pustaka Pelajar.

Boopathiraj, C & Chellamani, K. 2013. Analysis of Test Item on Difficulty Level and Discrimination Index in the Test For Research in Education. *International Journal of Social Science*. Vol. 2 No. 2.

Clegg, V. L., & Cshin, W. E. 1986. Improving Multiple-Choice Test. *Idea Paper Journal*. No 16

Depdiknas, 2006. *Permendiknas No 22 dan 23 tentang Standar Isi dan Kompetensi Lulusan untuk Satuan Pendidikan Dasar dan Menengah*. Jakarta: Depdiknas.

Djaali & Muljono, P. 2008. *Pengukuran Dalam Bidang Pendidikan*. Jakarta: Gramedia Widiasarana Indonesia.

Fitzpatric, A. 1983. The Meaning of Content Validity. *Applied Psychological Measurement*. Vol 7 No 1.

Kusaeri & Suprananto. 2012. *Pengukuran dan Penilaian Pendidikan*. Yokyakarta: Graha Ilmu.

Mansyur., Rasyid, H., & Suratno. 2009. *Asesmen Pembelajaran di Sekolah.* Yogyakarta: Multi Pressindo.

Mardapi, D. 2008. *Teknik Penyusunan Instrumen Tes dan Nontes*. Yogyakarta: Mitra Cendekia Press.

Mardapi, D.2012. *Pengukuran, Penilaian dan Evaluasi Pendidikan*. Yokyakarta: Nuha Medika.

Muslich, M. 2010. *Authentic Assessment: Penilaian Berbasis Kelas dan Kompetensi*. Bandung: Refika Aditama.

Naga, D. S. 1992. *Pengantar Teori Sekor Pada Pengukuran Pendidikan.*Jakarta: Gunadarma.

Purwanto. 2011. *Evaluasi Hasil Belajar*. Yokyakarta: Pustaka Pelajar.

Rasyid, H.& Mansur. 2007. *Penilaian Hasil Belajar*. Bandung: Wacana Prima.

Robert, T. 2006. The Use of Multiple Choice Tests for Formative and Summative Assessment. *Australian Computer Society, Inc*. Vol 52.

Ruslan. 2009.Validitas Isi. *Buletin Pa’ Biritta* No. 10. Tahun VI September 2009.

Subagyo,dkk. 2008. *Pendidikan Kewarganegaraan*. Semarang: UPT UNNES Pres

Sudaryono. 2012. *Dasar-dasar Evaluasi Pembelajaran*. Yogyakarta: Graha Ilmu.

Sudijono, A. 2011. *Pengantar Evaluasi Pendidikan*. Jakarta: Raja Grafindo Persada.

Sudjana, N. 2012. *Penilaian Hasil Proses Beljar Mengajar*. Bandung: Remaja Rosdakarya.

Sukardi, M. 2011. *Evaluasi Pendidikan Prinsip dan Operasionalnya*. Jakarta: Bumi Aksara.

Surapranata, S. 2004. *Panduan Penulisan Tes Tertulis.* Bandung: Remaja Rosdakarya.

Suryabrata, S. 2005. *Pengembangan Alat Ukur Psikologis*. Yogyakarta: Andi.

Susetyo. 2011. *Menyusun Tes Hasil Belajar Dengan Teori Ujian Klasik dan Responsi Butir*. Bandung: Cakra.

Trianto, 2010. *Model Pengembangan Terpadu dalam Teori dan Praktek. Prestasi*. Jakarta: Pustaka Publisher.

Widoyoko, E. P. 2012. *Teknik Penyusunan Instrumen Penelitian*. Yogyakarta: Pustaka pelajar.

Wirawan. 2011. *Evaluasi: Teori, Model, Standar, Aplikasi, dan Profesi.* Jakarta: Raja Grafindo Persada.

Ziegler, J & Detje, F. 2012. Aplication of Empirical Methodology to Evaluate Information Fusion Aproaches. *International Journal Methodology*. Vol 20.