

Cara Guru Dalam Mengembangkan Butir Tes Pilihan Ganda Sebagai Instrumen Evaluasi Hasil Belajar

Gusnita Efrina¹, Ayu Rahma Nengsi², Yona Syaida Oktira³
Universitas Nahdatul Ulama Sumatera Barat^{1,3}, IAIN Takengon²

ABSTRACT: *The low quality of teacher-made tests makes the information provided by the measurement results of these tests doubtful as a basis for educational decision making. In order to clearly identify the cause of the low quality of the test, it is important to conduct research that aims to look deeper into the methods that teachers take in the process of developing tests in the field. The sample was taken purposively, limited to 4 teachers at six grade, who were involved in the UTS and UAS question preparation team in Lubuk Alung District. The triangulation technique was chosen for data collection, the four samples were observed several times until the question writing process was completed. Interviews were conducted on matters that were not clear to the researcher during the process. Photos, videos and the form of test instruments made required was collected as documentations. The research findings show that basically the teacher has followed a part of the flow, although not all of the steps were followed correctly during the test item writing process. The findings of the study indicate the process of writing the items by the teacher, namely 1) the design of the lattice table, 2) the writing of the items was found: a) the choice of the operational words only measured aspects of knowledge and understanding, b) there were many negative types of questions such as the use of the word 'except' , c) the distractor is made very predictable, d) there is an ambiguous question language 3) the test review is carried out by a special team who edits all tests before they are printed and distributed, 4) testing and analysis of grain quality is not carried out at all, due to limitations funds, time and effort.*

Keywords: *The way the teacher develops the test, multiple choice test items, learning outcome evaluation instruments*

I. PENDAHULUAN

Banyak hasil penelitian yang menjelaskan lemahnya daya ukur dari tes pilihan ganda buatan guru. Soal yang tidak dikonstruksi dengan baik, banyaknya butir tes yang merupakan duplikat dari soal lama, dan tidak adanya review maupun revisi¹. Karakteristik butir dari instrumen tes yang dirancang juga tidak diketahui dengan jelas. Tidak adanya kesesuaian antara aspek yang diukur dengan tujuan pembelajaran yang dirumuskan atau proses pembelajaran yang terjadi¹. Rendahnya kualitas tes tersebut sering dipermasalahakan, disaat hasil pengukurannya dijadikan sebagai dasar pengambilan

¹ Wiguna, satria., syaukani.,ananda, rusydi. 2018. Kemampuan guru PAI dalam merancang tes. *Edu riliga*. vol.2 no 1.

keputusan Pendidikan². Minimnya skill juga sering disebut sebagai salah satu penyebab utama. Sehingga cara yang ditempuh oleh guru selama proses pengembangan tes, menjadi penting untuk diketahui lebih jauh, guna mengetahui secara lebih mendalam, kelemahan yang menjadi penyebab rendahnya kualitas tes tersebut.

Berbagai pelatihan dilakukan untuk meningkatkan kemampuan guru dalam merancang tes prestasi belajar³, namun sepertinya belum memberikan perubahan yang signifikan terhadap perbaikan kualitas tes yang dibuat. Salah satu cara yang mungkin dapat dilakukan adalah dengan melakukan penelusuran lebih dalam terkait pelaksanaan teknis perancangan tes ini di lapangan. Penelitian ini akan memberikan informasi terkait cara yang mereka tempuh dalam proses pengembangan, demi mendapatkan gambaran proses yang terjadi di lapangan. Dengan demikian akan diketahui setiap detail langkah yang dilakukan guru dalam mengembangkan butir tes prestasi belajar siswa. Sehingga akan semakin jelas diketahui tahapan pelaksanaan pengembangan tes yang keliru untuk diperbaiki.

Tes yang dikembangkan dengan mengikuti langkah-langkah penyusunan tes yang layak, akan menghasilkan butir tes berkualitas baik⁴. Penting sekali guru sebagai evaluator terdepan mengikuti secara benar setiap proses yang selayaknya harus ditepuh demi bisa menghasilkan tes yang melakukan fungsinya dengan baik. Dengan demikian tes tersebut akan mampu memberikan hasil penilaian yang benar terkait hal apa saja yang telah dipelajari dan tidak dipelajari oleh siswa dalam proses pembelajaran.

Berdasarkan uraian di atas, peneliti tertarik untuk mengetahui lebih jauh dengan cara mendeskripsikan secara mendalam cara yang ditempuh oleh guru, selama proses pengembangan tes pilihan ganda sebagai instrument.

II. METODE PENELITIAN

Penelitian ini menggunakan pendekatan kualitatif tujuannya untuk menjelaskan secara lebih dalam fakta lapangan terkait cara yang ditempuh oleh guru dalam mengembangkan tes prestasi belajar. Dalam penelitian ini peneliti merupakan instrumen kunci yang akan mengumpulkan data lapangan.

Prosedur penelitian dimulai dari observasi umum dengan mengikuti kegiatan rutin kelompok kerja guru sekolah dasar, khusus kelas VI Pada Kecamatan Lubuk Alung, Kabupaten Padang Pariaman sebanyak 2 kali. Kemudian dilakukan analisis data sekunder sebagai pembandingan dan memperjelas informasi dari hasil grand tour observasi. Hasil

² Suyata. 2010. Identifikasi Need Assessment: Studi Awal Model Pengembangan Bank Soal Berbasis Guru Di Provinsi DIY. *Jurnal Pendidikan*. Mei; 40 (1): 45-58

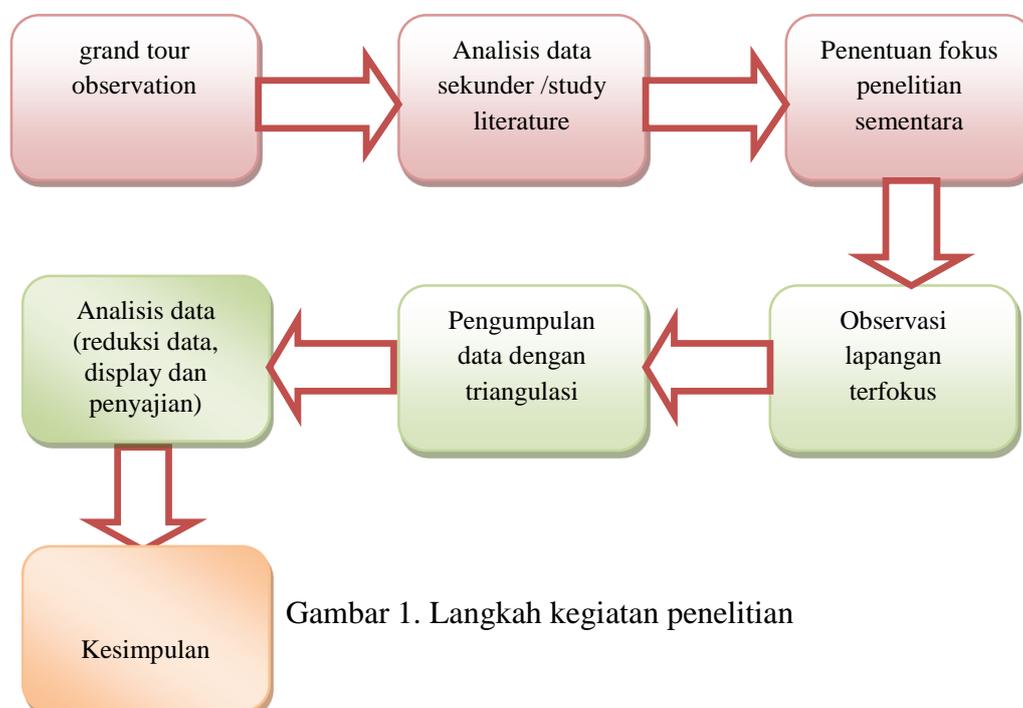
³ Osnal, Suhartoni, Imam Wahyudi. 2016. Meningkatkan Kemampuan Guru Dalam Menyusun Tes Hasil Belajar Melalui Workshop. *Jurnal Pancaran*. V.5 (1) hal 67-82

⁴ Azwar, S. 2010. Tes prestasi Fungsi Dan Pengembangan Pengukuran Prestasi Belajar. Jakarta: Pustaka Pelajar.

analisis tersebut dijadikan sebagai informasi dasar untuk memutuskan focus penelitian sementara.⁵ Setelah focus penelitian sementara ditentukan, dilanjutkan dengan observasi lapangan terfokus pada kegiatan yang sesuai dengan focus penelitian sementara, dan makin jelaslah focus penelitian yaitu pada cara yang ditempuh oleh guru dalam pengembangan tes pilihan ganda.

Selanjutnya yaitu pengumpulan data lapangan menggunakan teknik triangulasi 1) wawancara, 2) observasi partisipan fasif dan 3) dokumentasi. Teknik ini dipilih sekaligus untuk uji kredibilitas data penelitian. Narasumber atau sampel penelitian, diambil secara *purposive sampling* yaitu khusus pada guru yang terlibat dalam pembuatan soal tengah dan akhir semester, yang akan diwawancarai serta diobservasi secara berkala. 2 kali pengamatan pada kegiatan rutin kelompok kerja guru, 4 kali pengamatan secara mandiri di kediaman masing-masing narasumber yang berjumlah 4 orang. Jadi observasi dilakukan sebanyak satu kali untuk setiap narasumber. Namun tidak tertutup kemungkinan, guru yang tidak ikut membuat soal UTS dan UAS juga akan diwawancarai untuk pengkayaan data jika mereka membuat soal ulangan harian mata pelajaran tertentu. Dokumentasi dari tes buatan guru juga dikumpulkan untuk memperjelas temuan dari observasi dan wawancara.

Data yang terkumpul dianalisis dengan menggunakan model Miles dan Huberman⁶ dengan tiga cara yaitu: 1) Reduksi data: yaitu data dikelompokkan pada poin-poin tertentu, 2) Penyajian data yaitu data dikategorikan dan ditemukan polanya, 3) kesimpulan: hasil penelitian berdasarkan fakta lapangan. Berikut langkah penelitian dijelaskan dalam gambar 1. langkah kegiatan penelitian.



Gambar 1. Langkah kegiatan penelitian

⁵ Sugiyono. 2010. Metode penelitian kualitatif, kuantitatif dan R & D. Bandung: ALFABETA

III. KAJIAN TEORI

A. Tes Prestasi Belajar

Tes merupakan rangkaian pertanyaan, tugas-tugas yang harus dijawab dengan tujuan mengukur aspek tertentu dari peserta tes. Lebih jelasnya tes merupakan instrumen ukur untuk mengukur atribut psikologi. Tes dikelompokkan pada empat hal yaitu: a) tes untuk mengukur IQ umum untuk kemampuan seseorang dalam mengerjakan tugas, b) tes untuk mengukur bakat khusus tertentu, c) tes untuk mengukur prestasi belajar seseorang, dan d) tes untuk mengukur *personal assesment* ⁶.

Salah satu upaya untuk meningkatkan mutu pendidikan yaitu dengan meningkatkan kualitas proses dan hasil belajar sebagai bagian dari peningkatan kualitas pendidikan dapat dilakukan melalui sistem penilaian. Dalam proses penilaian proses dan hasil belajar siswa, aspek-aspek yang berkenaan dengan pemilihan jenis perangkat penilaian yang sesuai, dengan sasaran atau target belajar yang diukur, sangat berpengaruh terhadap kualitas kelulusan ⁷.

Perangkat instrumen yang baik akan memberikan informasi yang benar tentang perkembangan proses pendidikan yang dialami oleh peserta didik. Instrumen yang sering digunakan untuk mengukur hasil belajar siswa yaitu tes prestasi belajar. Secara umum instrumen tes prestasi yang baik penting untuk memperhatikan tingkat kesukaran, daya pembeda dan keberfungsian distraktor ⁸ serta memenuhi tuntutan validitas dan tuntutan reliabilitas ⁹, disamping itu, penting sekali memperhatikan penyusunan perangkat yang sesuai dengan prinsip-prinsip dan prosedur evaluasi yang layak, akan sangat menentukan validitas isi dari perangkat tes yang dikembangkan.

Secara umum instrumen tes prestasi yang baik penting untuk memperhatikan tingkat kesukaran, daya pembeda dan keberfungsian distraktor ⁹ serta memenuhi tuntutan validitas dan tuntutan reliabilitas ¹⁰

1. Taraf Sukar butir tes: Salah satu ciri butir soal yang baik adalah soal tidak terlalu sulit dan tidak terlalu mudah dimana tingkat kesukaran soal dipandang dari tingkat kesanggupan siswa dalam menjawabnya, bukan dilihat dari sudut guru sebagai pembuat soal ¹⁰

⁶ Poerwanti dan Masduki. 2009. Mengembangkan Tes Sebagai Instrumen Evaluasi. Chapter buku assesmen pembelajaran.

⁷ Ida, Farida. 2017. Evaluasi Pembelajaran Berdasarkan Kurikulum Nasional. Bandung : Rosdakarya

⁸ Arifin, Z. 2017. Kriteria Instrument Dalam Suatu Penelitian. *Jurnal THEOREMS*. Juli; 2 (1): 28-36

⁹ Gronlund, N.E., Linn, R.L., & Miller, M.D. 2009. Measurement & evaluation in teaching. Tenth edition. New York: Macmillan Publishing Co, Inc.

¹⁰ Djemari Mardapi 2012. Pengukuran, Penilaian, dan Evaluasi Pendidikan. Yogyakarta: Nuha Medika

2. Daya Beda butir tes: Daya beda soal berarti apakah soal yang dibuat mampu membedakan tingkat pemahaman antara siswa yang berkemampuan tinggi dengan siswa yang berkemampuan rendah ¹¹.
3. Distraktor: Jawaban yang salah yang sengaja di buat pada soal bentuk pilihan ganda disebut sebagai distraktor, tujuan pemasangan distraktor pada setiap butir tes yaitu agar dapat menjadi pilihan jawaban bagi peserta tes, dimana butir soal yang baik pengecohnya akan dipilih secara merata oleh peserta tes ¹². Distraktor telah dapat menjalankan fungsinya dengan baik apabila distraktor tersebut telah dipilih setidaknya 5% dari seluruh peserta tes

B. Tes Prestasi Buatan Guru

Otonomi daerah saat ini mengharuskan guru sebagai ganda terdepan dalam pendidikan, mengembangkan instrumen tes guna mengetahui perkembangan proses pembelajaran yang dialami siswa ³. Para Guru dalam praktik disekolah kecenderungannya tidak mampu mengembangkan tes prestasi berkualitas. Temuan penelitian menjelaskan banyak alasan menjadi penyebab perangkat tes buatan guru tidak mampu memberikan informasi yang akurat dalam melakukan pengukuran diantaranya: a) guru tidak selalu memahami arti pentingnya evaluasi yang tepat, b) guru tidak faham metode dan analisis butir tes, c) kegiatan analisis tes menyita banyak waktu. Hal ini menunjukkan kalau tes buatan guru tidak efektif digunakan sebagai alat untuk mengukur evaluasi hasil belajar siswa. pada akhirnya tes prestasi buatan guru belum tepat digunakan sebagai alat untuk mengukur hasil belajar karena tidak mampu memberikan informasi yang tepat yang bisa dijadikan landasan dalam pengambilan keputusan yang benar ¹³.

Dalam mengembangkan perangkat tes prestasi belajar, ketidakmampuan guru yang disebabkan oleh faktor eksternal karena kurangnya mendapat pelatihan menyusun tes, juga disebabkan oleh faktor internal dari para guru itu sendiri yaitu: a) guru kesulitan menilai kualitas tes, b) kurang mengerti dalam menentukan jumlah butir soal dengan waktu yang tersedia, c) kebingungan dalam membuat pedoman penskoran ¹.

IV. HASIL PENELITIAN

Hasil penelitian menunjukkan jenis tes yang paling sering digunakan disekolah untuk mengukur proses pembelajaran, adalah tes bentuk pilihan ganda. Tes ini dipilih karena dinilai mudah dalam proses penggunaan dan pemeriksaan serta mampu mengukur secara komprehensif materi yang telah diberikan dalam kurun waktu tertentu. Giel

¹¹ Sufanti, M. 2013. Tingkat Kesulitan dan Daya Beda Soal Tes Sumatif Bahasa Indonesia Berdasarkan Kurikulum 2013 SMP. *Jurnal Kajian Linguistic Dan Sastra*. Hal 26

¹² Sadtyadi, H & Kartowagiran, B. 2014. Pengembangan Instrumen Penilaian Kinerja Guru Sekolah Dasar Berbasis Tugas Pokok dan Fungsi. *Jurnal Penelitian dan Evaluasi Pendidikan*. 18 (2) : 290-304

³ Suyata. 2010. Identifikasi Need Assessment: Studi Awal Model Pengembangan Bank Soal Berbasis Guru Di Provinsi DIY. *Jurnal Pendidikan*. Mei; 40 (1): 45-58

¹³ Sudijono, anas. 2012. Pengantar evalausi pendidikan. Jakarta: PT Raja Grafindo Persada

menjelaskan tes bentuk pilihan ganda merupakan salah satu instrumen penilaian yang paling efektif digunakan dalam evaluasi pendidikan hingga sekarang ¹⁴.

Tes yang baik dan valid dalam mengukur kemampuan siswa akan memberikan informasi yang tepat terhadap evaluasi belajar siswa. Tapi tes buatan guru banyak yang tidak dapat dipertanggung jawabkan hasil pengukurannya, karena kemampuan daya ukur yang memang masih minim ³. Selain kemampuan guru yang minim, cara mereka dalam merancang tes patut ditelusuri lebih jauh, untuk mengetahui secara jelas teknis perancangan tes dilapangan.

Tes pilihan ganda yang dirancang oleh guru ditujukan untuk mengukur hasil belajar siswa setelah melewati serangkaian pembelajaran. Kegiatan pengembangan tes dibuat untuk ulangan harian, ujian tengah semester dan ujian akhir semester. Dalam penelitian ini akan dibahas beberapa kegiatan penulisan butir tes pilihan ganda yang dilakukan oleh guru dalam proses pengembangannya. Kegiatan yang teramati dapat dikelompokkan pada tiga kegiatan utama yaitu: perancangan tabel kisi-kisi, penulisan butir tes, review validitas isi. Dari tiga kegiatan yang dikelompokkan, observasi lebih difokuskan pada cara atau teknik yang dipakai oleh guru ketika melakukan kegiatan penulisan butir tes.

1. Perancangan tabel kisi-kisi tes

Kegiatan perancangan tabel kisi-kisi dilakukan oleh semua narasumber yang diamati. Tabel kisi-kisi yang dibuat berisikan beberapa aspek yaitu standar kompetensi, kompetensi dasar, indikator, materi pembelajaran, nomor butir soal, dan jumlah butir soal. Berikut ditunjukkan dalam tabel 4.1 kisi-kisi tes buatan guru

Tabel 4.1 Kisi-kisi Mata Pelajaran.....

Kompetensi dasar	Materi	Indikator	No Soal	Jumlah Soal

Sumber: dokumentasi sekolah

2. Penulisan butir tes

Pada kegiatan penulisan butir soal, pengamatan difokuskan pada penulisan tes bentuk pilihan ganda. Hasil pengamatan dapat diklasifikasikan pada beberapa hal sebagai berikut:

a) Jumlah dan sebaran butir

Perangkat tes ujian tengah dan akhir semester yang dibuat oleh guru terdiri atas 40 soal, dengan penyebaran 35 butir soal bentuk pilihan ganda dan 5 butir soal uraian.

¹⁴ Mark J.Gierl, Okan Bulut, Qi Guo, And Xinxing Zhang. 2017. Developing, Analyzing, And Using Distractors For Multiple-Choice Test In Education: A Comprehensive Review. *Review Of Educational Research*. V.87 (6) pp.1082-1116. DOI:10.302/0034654317726529

Observasi dilakukan sebanyak 6 kali. 2 kali pada saat narasumber mengikuti kegiatan kelompok kerja guru di gugus wilayah 1 Kecamatan Lubuk Alung, Kabupaten Padang Pariaman, dan 4 kali dilakukan pada narasumber secara pribadi di kediaman masing-masing.

b) Pilihan kata-kata operasional

Pada saat penulisan butir tes, narasumber membuat soal berdasarkan indikator yang disusun dalam tabel kisi-kisi. Soal yang dibuat kebanyakan mengukur kemampuan kognitif pengetahuan dan pemahaman dalam tingkatan taksonomi bloom. Sedangkan dalam tabel kisi-kisi tidak dituliskan tingkatan kemampuan kognitif yang ingin diukur, namun dari kata-kata operasional yang digunakan dalam soal jelas terlihat. Kata operasional yang sering digunakan dalam soal yang dibuat seperti “apakah..., sebutkanlah..., siapakah..., jelaskanlah..., mengapa...” pemakaian kata-kata operasional tersebut sering muncul dalam butir soal yang ditulis. Berikut contoh soal yang dibuat dalam mata pelajaran ilmu pengetahuan sosial kelas VI sekolah dasar

“Siapakah pahlawan nasional berikut yang berasal dari provinsi aceh? “

“apakah yang dimaksud dengan kegiatan produksi?”

“sebutkanlah 2 bentuk peninggalan bukti keberadaan kerajaan islam dinusantara?”

Dari 35 soal bentuk pilihan ganda yang ditulis untuk mata pelajaran IPS, terdapat 16 soal yang menggunakan atau merujuk pada penggunaan kata-kata operasional tersebut. Ketika ditanya terkait kemampuan kognitif siswa diukur hingga tingkatan berapa dalam taksonomi bloom?, keempat narasumber memberikan jawaban yang kurang lebih sama. Untuk kelas VI soal yang dibuat hendaknya dapat mengukur kognitif siswa pada tingkatan C1- C4, sesuai dengan ketentuan standar evaluasi dalam kurikulum, begitu pemahaman mereka. Namun dalam pelaksanaan narasumber terlihat kurang faham membuat butir soal yang sesuai untuk mengukur kemampuan hingga tingkatan C4, kebanyakan soal yang dibuat hanya untuk mengukur kemampuan pengetahuan dan pemahaman.

c) Pembuatan option jawaban/ distraktor

Soal bentuk pilihan ganda memiliki beberapa *option* atau pilihan jawaban yang mana hanya ada satu pilihan jawaban yang benar, sisanya merupakan jawaban yang salah salah. Jawaban yang salah merupakan pengecoh yang sengaja dibuat, untuk membuat para peserta tes memilih jawaban yang benar, dari apa yang telah dipelajari dari beberapa pilihan jawaban yang berbeda yang disediakan oleh soal.

Pembuatan pilihan jawaban untuk soal yang dibuat oleh narasumber diamati, terlihat tidak semua narasumber membuat pengecoh untuk soal-soal tersebut dengan cara yang benar. Beberapa diantara mereka bahkan membuat option jawaban salahnya sangat mudah dan jauh menyimpang dari jawaban yang benar. Sehingga berkemungkinan akan sangat mudah ditebak oleh para peserta tes jawaban yang benarnya. Persoalan seperti Ini

terlihat pada lebih dari 10 soal pilihan ganda yang dibuat, hal terlihat pada semua narasumber yang diamati.

Ketika ditanyai kenapa membuat pengecohnya mudah sekali ditebak oleh siswa, jawaban mereka kurang lebih merujuk pada hal yang sama. Seperti agar siswa mudah mengenali jawaban yang benarnya, dan akan banyak siswa yang menjawab dengan benar soal tersebut. Ada juga yang menjawab kalau pilihan jawaban salahnya terlalu dekat persamaannya dengan jawaban yang benarnya, akan menjadikan soal tersebut menjadi soal yang sangat sulit bagi siswa.

d) Pilihan bahasa

Terkait pemakaian bahasa dalam penulisan butir soal pilihan ganda, narasumber membuat soal menggunakan bahasa Indonesia yang baik serta mudah difahami, tidak ada terlihat butir soal yang menggunakan bahasa daerah ataupun bahasa asing. Dan juga ditemukan penulisan soal dengan tipe negative. Dari 35 buah butir pilihan ganda yang dibuat, terlihat 5 butir diantaranya soal dengan tipe negatif, contohnya: 1) “Gambar dibawah ini yang tergolong kegiatan konsumsi adalah kecuali...?”, 2) gambar dibawah ini yang bukan merupakan kegiatan produksi adalah...?.

e) Pengambilan soal lama

Dalam kegiatan penulisan butir soal, selain menulis soal baru yang sesuai dengan indikator yang telah dirumuskan, juga terlihat aktifitas dimana narasumber yang diamati mengambil soal-soal lama dari dokumentasi sekolah, kemudian dijadikan salah satu soal dalam perangkat tes yang sedang ditulis. 4 narasumber yang diamati, semuanya melakukan hal yang sama, ketika mereka sudah mentok dan tidak punya ide lagi saat menulis soal yang sesuai dengan indikator. Soal lama yang diambil untuk perangkat tes yang sedang ditulis, disesuaikan dengan indikator yang telah disusun dalam tabel kisi-kisi yang dirancang. Namun ada juga 1 orang narasumber terlihat mengambil soal lama untuk dijadikan bagian dalam tes yang sedang dirancang, namun tidak mengambil soal yang sesuai dengan indikator, soal diambil secara random dan dimasukkan secara acak juga dalam perangkat tes yang sedang disusun.

f) Pembuatan kunci jawaban dan pedoman penskoran

Setelah jumlah soal yang ditulis telah sesuai dengan jumlah yang ditentukan dalam tabel kisi-kisi, selanjutnya narasumber membuat kunci jawab dari semua soal tersebut, sekaligus membuat pedoman penskoran. Hasil pengamatan menunjukkan semua narasumber membuat kunci jawaban dan pedoman penskoran dari perangkat tes yang mereka buat.

3. Review tes

Kegiatan terakhir dari proses pembuatan soal ujian adalah pembuatan kunci jawaban dan pedoman penskoran. Ini adalah jawaban dari keempat narasumber yang diamati secara intensif selama proses penelitian. Terlihat juga keempat narasumber

tersebut menyerahkan tes yang telah mereka tulis kepada ketua gugus kelompok kerja yang membawahi mereka. Hal ini dilakukan untuk dilakukan proses review, dimana ketua gugus telah membentuk tim review yang terdiri dari dua orang. Tugasnya adalah membaca ulang soal-soal yang telah dirancang oleh guru yang ditunjuk sebagai pembuat soal mata pelajaran tertentu.

Tim review soal yang diwawacari menyatakan soal yang mereka terima dibaca ulang, untuk memeriksa dengan banar kesalahan pengetikan, merevisi bahasa soal yang dianggap rancu dan melakukan pengeditan teknis penulisan soal untuk disesuaikan dengan templete yang telah ditentukan. Tahap review ini adalah tahap finishing dari pembuatan soal sebelum soal dicetak dan didistribusikan.

V. SIMPULAN

Berdasarkan uraian dari temuan hasil penelitian, dapat disimpulkan sebagai berikut:

1. Secara umum langkah yang diikuti oleh guru dalam perancangan tes bentuk pilihan ganda terlihat mengikuti alur, namun langkah tersebut belum selesai, karena tidak dilakukan ujicoba tes guna mengetahui kualitas butir dari tes.
2. Dalam penulisan butir tes banyak soal yang diambil dari soal-soal lama, yang dirasa sesuai dengan indikator, namun terlihat juga soal yang diambil tidak sesuai dengan kisi-kisi yang ada.
3. Pada penulisan butir tes terlihat guru tidak memahami dengan benar cara memilih kata-kata operasional yang dapat mengukur kognitif siswa sesuai dengan tingkatan kemampuan yang ingin diukur (taksonomi bloom) dan banyak soal yang berjenis negative menggunakan kata “kecuali”.

DAFTAR PUSTAKA

- Arifin, Z. 2017. Kriteria Instrument Dalam Suatu Penelitian. *Jurnal THEOREMS*. Juli; 2 (1): 28-36
- Azwar, S. Tes prestasi Fungsi Dan Pengembangan Pengukuran Prestasi Belajar. Jakarta: Pustaka Pelajar. 2010
- Djemari Mardapi 2012. Pengukuran, Penilaian, dan Evaluasi Pendidikan. Yogyakarta: Nuha Medika

- Gronlund, N.E., Linn, R.L., & Miller, M.D. 2009. *Measurement & evaluation in teaching*. Tenth edition. New York: Macmillan Publishing Co, Inc.
- Ida, Farida. 2017. *Evaluasi Pembelajaran Berdasarkan Kurikulum Nasional*. Bandung : Rosdakarya
- Mark J.Gierl, Okan Bulut, Qi Guo, And Xinxing Zhang. 2017. Developing, Analyzing, And Using Distractors For Multiple-Choice Test In Education: A Comprehensive Review. *Review Of Educational Research*. V.87 (6) pp.1082-1116. DOI:10.302/0034654317726529
- Osnal, Suhartoni, Imam Wahyudi. 2016. Meningkatkan Kemampuan Guru Dalam Menyusun Tes Hasil Belajar Melalui Workshop. *Jurnal Pancaran*. V.5 (1) hal 67-82
- Poerwanti dan Masduki. 2009. *Mengembangkan Tes Sebagai Intrumen Evaluasi*. Chapter buku assesmen pembelajaran.
- Sufanti, M. 2013. Tingkat Kesulitan dan Daya Beda Soal Tes Sumatif Bahasa Indonesia Berdasarkan Kurikulum 2013 SMP. *Jurnal Kajian Linguistic Dan Sastra*. Hal 26
- Sadtyadi, H & Kartowagiran, B. 2014. Pengembangan Intrumen Penilaian Kinerja Guru Sekolah Dasar Berbasis Tugas Pokok dan Fungsi. *Jurnal Penelitian dan Evaluasi Pendidikan*. 18 (2) : 290-304
- Suyata. 2010. Identifikasi Need Assessment: Studi Awal Model Pengembangan Bank Soal Berbasis Guru Di Provinsi DIY. *Jurnal Pendidikan*. Mei; 40 (1): 45-58
- Sudijono, anas. 2012. *Pengantar evalausi pendidikan*. Jakarta: PT Raja Grafindo Persada
- Sugiyono. 2010. *Metode penelitian kualitatif, kuantitatif dan R & D*. Bandung: ALFABETA
- Wiguna, satria., syaukani.,ananda, rusydi. 2018. Kemampuan guru PAI dalam merancang tes. *Edu riliga*. vol.2 no 1.
- Zamsir. 2010. Kualitas tes buatan guru matematika SLTP negeri di kota madya kendari. *Jurnal pendidikan matematika dan sains*. Vol.2. 65-69