

BAB II

KAJIAN PUSTAKA

A. Penelitian yang Relevan

1. Penelitian yang dilakukan oleh Andi Surahma Halik (UIN Alauddin Makassar) pada tahun 2017 yang berjudul “Analisis Butir Soal Ujian Akhir Sekolah (UAS) Mata Pelajaran Matematika Pada Tahun Ajaran 2015/2016 SMP Negeri 36 Makassar”. Dalam penelitian tersebut dapat diketahui hasil penelitian tersebut yaitu:
 - a) Berdasarkan kriteria tingkat kesukaran soal diperoleh soal sangat mudah 2 butir (5%), sedang 24 butir (60%), sukar 12 butir (30%), dan sangat sukar 2 butir (5%).
 - b) Berdasarkan kriteria daya beda soal diperoleh soal dengan daya beda sangat jelek 3 butir (7,5%), jelek 12 butir (30%), cukup 14 butir (35%), dan baik 11 butir (27,5%).
 - c) Berdasarkan keefektifan pengecoh diperoleh 28 butir (70%) dalam kategori sangat baik, 9 butir (22,5%) dalam kategori baik, dan 3 butir (7,5%) dalam kategori buruk.

Persamaan dengan penelitian yang dilakukan oleh penulis yaitu sama-sama menganalisis tentang kualitas butir soal. teknik pengumpulan data yaitu dokumentasi dan dianalisis dengan bantuan aplikasi Anates Versi 4. Sedangkan perbedaannya terletak pada soal yang dianalisis, skripsi yang dilakukan oleh Andi Surahma Malik menganalisis soal Ujian Akhir Sekolah mata pelajaran Matematika tingkat SMP. Sedangkan penulis menganalisis soal Penilaian Harian

Bersama II mata pelajaran Bahasa Indonesia tingkat SMA. Analisis yang dilakukan oleh Andi Surahma Malik hanya dari segi tingkat kesukaran, daya beda dan efektivitas pengecoh. Sedangkan penulis menganalisis dari segi validitas, reliabilitas, daya pembeda, tingkat kesukaran dan efektivitas pengecoh.

2. Penelitian yang dilakukan oleh Reni Setianingsih (IAIN Purwokerto) pada tahun 2018 yang berjudul “Analisis Kualitas Butir Soal Ujian Akhir Semester 2 Mata Pelajaran Sejarah Kebudayaan Islam Kelas VIII Di MTs Ma’arif Nu Kemranjen Kabupaten Banyumas Tahun Ajaran 2016/2017”. Dalam penelitian tersebut dapat diketahui hasil penelitian yaitu:

50 butir soal ujian akhir semester 2 mata pelajaran Sejarah Kebudayaan Islam yang diikuti oleh 191 siswa di 3 MTs Ma’arif NU Kemranjen Kabupaten Banyumas belum baik. Berdasarkan analisis yang dilakukan dari Tingkat kesukaran, daya pembeda dan fungsi distraktor dapat diketahui: tingkat kesukaran meliputi kategori sangat sukar 2 soal, sukar 5 soal, sedang 33 soal, mudah 5 soal, dan sangat mudah 5 soal. Daya pembeda meliputi: sangat baik 4 soal, baik 22 soal, cukup 12 soal, jelek 10 soal, dan sangat jelek 2 soal. Dilihat dari fungsi pengecoh ditemukan Fungsi Distraktor dengan kriteria berfungsi sebanyak 33. Terdapat 13 butir soal yang hanya dua pengecohnya berfungsi dengan baik. Terdapat 2 butir soal yang hanya satu pengecohnya berfungsi. Terdapat 2 butir soal yang fungsi distraktornya tidak berfungsi.

Persamaan dengan penelitian yang dilakukan oleh penulis yaitu sama-sama menganalisis tentang kualitas butir soal. Sedangkan perbedaannya terletak

pada teknik pengumpulan data, skripsi yang ditulis oleh Reni Setianingsih menggunakan teknik observasi, wawancara, dan dokumentasi. Sedangkan teknik pengumpulan data yang dilakukan penulis hanya teknik dokumentasi. Soal yang dianalisis oleh Reni Setianingsih yaitu soal Ujian Akhir Semester 2 mata pelajaran Sejarah Kebudayaan Islam tingkat MTs. Sedangkan penulis menganalisis soal Penilaian Harian Bersama II mata pelajaran Bahasa Indonesia tingkat SMA.

B. Landasan Teori

1. Penilaian Hasil Belajar

a. Pengertian Penilaian Hasil Belajar

Daryanto (2014:111) mengungkapkan bahwa penilaian adalah rangkaian kegiatan untuk memperoleh, menganalisis, dan menafsirkan data tentang proses dan hasil belajar peserta didik yang dilakukan secara sistematis dan berkesinambungan sehingga dapat menjadi informasi yang bermakna dalam pengambilan keputusan. Diungkapkan oleh Mundilarto (2012:14) bahwa proses penilaian (*assessment*) mencakup pengumpulan bukti-bukti atau informasi yang menunjukkan tingkat pencapaian belajar peserta didik. Sudjana (2012:22-23) mengemukakan bahwa hasil belajar adalah kemampuan-kemampuan yang dimiliki siswa setelah mengikuti pembelajaran. Kemampuan-kemampuan yang dimiliki siswa baik yang bersifat kognitif, afektif maupun psikomotor yang diperoleh setelah peserta didik mengikuti proses belajar mengajar. Berdasarkan pendapat-pendapat diatas maka dapat disimpulkan bahwa penilaian hasil belajar yaitu rangkaian kegiatan mengumpulkan bukti-bukti atau informasi tentang proses dan hasil belajar peserta didik setelah mengikuti pembelajaran yang dilakukan

secara sistematis dan berkesinambungan sehingga dapat digunakan untuk pengambilan keputusan.

b. Tujuan Penilaian Hasil Belajar

Tujuan penilaian hasil belajar peserta didik menurut Kunandar (2013:70) adalah sebagai berikut:

1) Melacak kemajuan peserta didik.

Melakukan penilaian akan menghasilkan perkembangan hasil peserta didik. Perkembangan hasil peserta didik dapat diidentifikasi yaitu meningkat atau menurun. Guru bisa membuat data kemajuan peserta didik yang berisi pencapaian hasil belajar peserta didik secara periodik.

2) Mengecek ketercapaian kompetensi peserta didik.

Melakukan penilaian akan dapat mengetahui apakah peserta didik telah menguasai kompetensi tersebut atau belum. Apabila masih ada peserta didik yang belum menguasai kompetensi tertentu maka akan dicarikan tindakan seperti remedial. Hal ini bertujuan untuk membantu peserta didik mencapai kompetensi tertentu.

3) Mendeteksi kompetensi yang belum dikuasai oleh peserta didik.

Adanya penilaian maka dapat mendeteksi kompetensi-kompetensi yang belum dikuasai oleh peserta didik. Hal ini ditujukan untuk memperoleh data tentang kompetensi yang belum dikuasai oleh peserta didik. Sehingga guru dapat mengulang kembali kompetensi tertentu yang belum dikuasai peserta didik.

4) Menjadi umpan balik untuk perbaiki bagi peserta didik.

Adanya penilaian maka guru dapat memperoleh bahan acuan untuk memperbaiki hasil belajar peserta didik yang masih dibawah kriteria ketuntasan minimum (KKM).

c. Prinsip dan Prosedur Penilaian Hasil Belajar

Pentingnya penilaian hasil belajar peserta didik dalam menentukan kualitas pendidikan, maka upaya merencanakan dan melaksanakan penilaian hendaknya memperhatikan prinsip dan prosedur penilaian. Sudjana (2012:8) mengemukakan prinsip dan prosedur penilaian hasil belajar sebagai berikut:

- 1) Menilai hasil belajar hendaknya dirancang sedemikian rupa, sehingga jelas abilitas yang harus dinilai, materi penilaian, alat penilaian, interpretasi penilaian. Patokan atau rambu-rambu dalam merancang penilaian hasil belajar adalah kurikulum yang berlaku dan buku pelajaran yang digunakan. Kurikulum dan buku yang digunakan hendaknya dipelajari tujuannya, pokok bahasan yang diberikan, ruang lingkup dan urutan penyajian, serta pedoman pelaksanaannya.
- 2) Penilaian hasil belajar hendaknya menjadi bagian integral dari proses belajar mengajar. Artinya, penilaian senantiasa dilaksanakan pada setiap saat proses belajar mengajar sehingga pelaksanaannya berkesinambungan.
- 3) Penilaian harus menggunakan berbagai alat penilaian dan sifatnya komprehensif. Sifat komprehensif dimaksudkan segi atau abilitas yang dinilai tidak hanya aspek kognitif, afektif dan aspek psikomotor, melainkan semua

aspek digunakan dalam penilaian. Hal ini dilakukan untuk menilai peserta didik secara objektif.

- 4) Penilaian hasil belajar hendaknya diikuti dengan tindakan lanjutan. Setelah melakukan penilaian guru sudah memiliki hasil belajar peserta didik. Oleh karena itu guru harus melakukan kegiatan lanjutan yaitu mengelompokkan peserta didik yang sudah mencapai kompetensi dan peserta didik yang belum mencapai kompetensi. Peserta didik yang sudah mencapai kompetensi bisa diberi tugas pengayaan untuk memperdalam materi dan peserta didik yang belum mencapai kompetensi dapat mengikuti remedial.

Dari pemaparan prinsip dan prosedur penilaian hasil belajar tersebut dapat disimpulkan bahwa merencanakan dan melaksanakan penilaian harus berpedoman pada kurikulum yang berlaku dan dilakukan secara berkesinambungan. Penilaian juga harus bersifat objektif kepada peserta didik serta guru harus memiliki tindakan lanjutan setelah melakukan penilaian hasil belajar.

2. Instrumen Penilaian

a. Pengertian Instrumen Penilaian

Trianto (2014:129) mengungkapkan bahwa instrumen merupakan alat yang dapat digunakan untuk mengukur tingkat ketercapaian kompetensi. Bentuk instrumen merupakan alat yang digunakan dalam melakukan penilaian/pengukuran/evaluasi terhadap pencapaian kompetensi peserta didik dalam bentuk tes maupun non tes. Penilaian menurut Trianto (2014:123) yaitu serangkaian kegiatan untuk memperoleh, menganalisis, dan menafsirkan data

tentang hasil belajar peserta didik yang dilakukan secara sistematis dan berkesinambungan sehingga menjadi informasi yang bermakna dalam pengambilan keputusan. Sedangkan Abidin (2014:66) mengungkapkan bahwa penilaian merupakan bagian dari kegiatan evaluasi yang berfokus pada dimensi pembelajaran yang di dalamnya terkandung juga istilah tes dan pengukuran. Dari beberapa pendapat tersebut dapat disimpulkan instrumen penilaian yaitu alat evaluasi yang digunakan untuk mengukur tingkat ketercapaian kompetensi peserta didik dalam bentuk tes maupun non tes yang dilakukan secara sistematis dan berkesinambungan.

b. Tes

1) Pengertian Tes

Menurut Sudaryono (2012: 101) tes merupakan himpunan pertanyaan yang harus dijawab, harus ditanggapi, atau tugas yang harus dilaksanakan oleh orang yang di tes. Tes digunakan untuk mengukur sejauh mana seorang siswa telah menguasai pelajaran yang disampaikan terutama meliputi aspek pengetahuan dan ketrampilan. Menurut Sudijono (2012: 67) tes merupakan alat atau prosedur yang digunakan dalam rangka mengukur dan menilai di bidang pendidikan yang berbentuk pemberian tugas sehingga dapat dihasilkan nilai yang menggambarkan tingkah laku, di mana nilai dapat dibandingkan dengan nilai standar tertentu. Dapat diambil kesimpulan bahwa tes merupakan suatu alat ukur untuk mengukur dan menilai yang berbentuk pemberian tugas, yang harus dijawab, harus ditanggapi, atau tugas yang harus dilaksanakan oleh orang yang di tes

2) Ciri-ciri Tes yang Baik

Menurut Arikunto (2013: 72) sebuah tes dikatakan baik sebagai alat pengukuran apabila memenuhi persyaratan sebagai berikut:

1) Validitas

Sebuah data dapat dikatakan valid apabila sesuai dengan keadaan senyatanya. Data yang dihasilkan oleh instrumen yang benar, valid dan sesuai dengan kenyataannya maka instrumen yang digunakan juga valid. Sebuah tes dikatakan valid atau tepat apabila tes tersebut dapat mengukur apa yang hendak diukur.

2) Reliabilitas

Tes dikatakan reliabel apabila memberikan hasil yang tetap walaupun diteskan berkali-kali. Dengan kata lain, yang dimaksud dari reliabilitas adalah ketetapan dari sebuah instrumen.

3) Objektivitas

Sebuah tes memiliki objektivitas apabila dalam melaksanakan tes tersebut tidak ada faktor subjektif yang memengaruhi.

4) Praktikabilitas

Sebuah tes dikatakan memiliki praktikabilitas yang tinggi apabila tes tersebut bersifat praktis. Tes yang praktis adalah tes yang mudah dalam pelaksanaannya, mudah pemeriksaannya, mudah dalam pengadministrasiannya, dan dilengkapi petunjuk yang jelas sehingga dapat diberikan oleh orang lain.

5) Ekonomis

Sebuah alat atau instrumen dikatakan ekonomis apabila dalam pelaksanaannya tidak memerlukan biaya yang mahal, waktu dan tenaga yang banyak.

3) Prinsip Dasar Penyusunan Tes

Dalam menyusun tes sebagai alat untuk mengukur kemampuan peserta didik perlu diperhatikan beberapa prinsip dasar penyusunan tes hasil belajar. Sudijono (2012: 97-99) memaparkan beberapa prinsip dasar penyusunan tes hasil belajar sebagai berikut:

- 1) Tes hasil belajar harus dapat mengukur hasil belajar dengan jelas sesuai tujuan instruksional. Adanya kejelasan tersebut mempermudah penyusun tes dalam menyusun butir-butir soal.
- 2) Butir-butir soal tes hasil belajar harus dapat mewakili materi pelajaran yang telah diajarkan.
- 3) Bentuk soal yang digunakan dalam tes hasil belajar harus bervariasi, sehingga tepat untuk mengukur hasil belajar sesuai dengan tujuan diadakan tes.
- 4) Desain tes hasil belajar harus disusun relevan dengan kegunaannya untuk memperoleh hasil yang diinginkan.
- 5) Tes hasil belajar harus memiliki tingkat reliabilitas yang baik dan dapat diandalkan. Apabila tes hasil belajar dilaksanakan berkali-kali terhadap subjek

yang sama pada waktu yang berbeda, hasilnya akan selalu sama atau relatif sama.

- 6) Selain untuk alat pengukur keberhasilan belajar peserta didik, tes hasil belajar juga harus dapat memberikan informasi yang dapat digunakan untuk memperbaiki cara belajar peserta didik dan cara mengajar guru.

3. Analisis Butir Soal

a. Pengertian Analisis Butir Soal

Menurut Nurgiyantoro (2016:214) analisis butir soal adalah identifikasi jawaban benar dan salah tiap butir soal yang diujikan oleh peserta didik. Arikunto (2013: 220) mengemukakan bahwa analisis butir soal adalah suatu prosedur yang sistematis, yang akan memberikan informasi-informasi yang sangat khusus terhadap butir tes yang kita susun. Selain itu, menurut Sudjana (2012: 135) analisis butir soal adalah pengkajian pertanyaan-pertanyaan tes agar diperoleh perangkat pertanyaan yang memiliki kualitas yang memadai.

Berdasarkan paparan beberapa ahli mengenai analisis butir soal, kesimpulannya bahwa analisis butir soal merupakan suatu kegiatan mengkaji dan mengidentifikasi setiap butir soal guna mengetahui kualitas setiap butir soal tersebut. Hasil dari proses mengkaji dan mengidentifikasi soal dapat digunakan untuk melakukan perbaikan dan/atau penyempurnaan pada setiap butir soal.

b. Manfaat Analisis Butir Soal

Kegiatan analisis pada setiap butir soal tentunya memberikan manfaat bagi penyusun soal. Manfaat tersebut menurut Kusaeri dan Suprananto (2012: 164) antara lain:

- 1) Menentukan soal-soal yang cacat atau tidak berfungsi dengan baik.
- 2) Meningkatkan butir soal melalui tiga komponen analisis yaitu tingkat kesukaran, daya pembeda, dan pengecoh soal.
- 3) Merevisi soal yang tidak relevan dengan materi yang diajarkan, ditandai dengan banyaknya anak yang tidak dapat menjawab butir soal tertentu.

c. Teknik Analisis Butir Soal

1. Validitas

Arifin (2014: 247) menyebutkan ada dua unsur penting dalam validitas yaitu validitas menunjukkan suatu derajat dan validitas selalu dihubungkan dengan suatu putusan atau tujuan yang spesifik. Suatu instrumen pengukur dapat dikatakan mempunyai validitas yang tinggi apabila instrumen tersebut menjalankan fungsi ukurnya, atau memberikan hasil ukur yang sesuai dengan maksud dilakukannya pengukuran tersebut. Selain itu, menurut Surapranata (2005: 50) validitas adalah suatu konsep yang berkaitan dengan sejauh mana tes yang telah mengukur apa yang seharusnya diukur. Menurut Sudijono (2012: 163) terdapat dua macam validitas yaitu validitas tes dan validitas item.

a) Validitas Tes

Validitas tes merupakan pengukuran yang digunakan untuk soal yang akan digunakan secara keseluruhan. Pengukuran validitas tes dapat dilakukan secara rasional dan secara empirik.

1) Validitas Rasional

Sudijono (2012: 164) memaparkan bahwa “validitas rasional adalah validitas yang diperoleh atas dasar hasil pemikiran, validitas yang diperoleh dengan berpikir secara logis”. Validitas rasional untuk sebuah instrumen evaluasi menunjuk pada kondisi sebuah instrumen yang memenuhi persyaratan valid berdasarkan penalaran. Jadi suatu tes dikatakan memiliki validitas rasional apabila setelah dilakukan analisis secara rasional menunjukkan bahwa tes hasil belajar tersebut memang dapat mengukur apa yang seharusnya diukur. Ada dua macam validitas rasional yang dapat dicapai sebuah instrumen yaitu validitas isi dan validitas konstruksi.

(a) Validitas isi

Sudijono (2012: 164) mengungkapkan bahwa validitas isi menunjuk pada suatu kondisi sebuah instrumen yang disusun berdasarkan isi materi pelajaran yang dievaluasi. Validitas isi merupakan validitas yang dilihat dari segi isi tes sebagai alat pengukur hasil belajar.

(b) Validitas konstruksi

Surapranata (2005: 53) mengungkapkan bahwa validitas konstruk menunjuk pada suatu kondisi sebuah instrumen yang disusun berdasarkan konstruk aspek-aspek kejiwaan yang seharusnya dievaluasi. Konstruk merupakan

sesuatu yang berkaitan dengan fenomena dan objek yang abstrak, tetapi gejalanya dapat diamati dan diukur.

2) Validitas Empirik

Sudijono (2012: 167) mengemukakan validitas empirik adalah validitas yang bersumber pada atau diperoleh atas dasar pengamatan dilapangan. Sebuah instrumen dikatakan memiliki validitas empirik apabila sudah diuji dari pengalaman atau pengamatan di lapangan, dan terbukti bahwa tes tersebut dapat mengukur apa yang seharusnya diukur. Validitas empirik dapat dilihat dari dua segi yaitu validitas ramalan dan validitas bandingan.

(a) Validitas ramalan

Menurut Surapranata (2005: 54) validitas prediksi menunjuk pada hubungan antara tes skor yang diperoleh peserta tes dengan keadaan yang akan terjadi diwaktu mendatang. Sebuah tes dikatakan memiliki validitas prediksi apabila mempunyai kemampuan untuk meramalkan yang akan terjadi pada masa mendatang. Untuk mengetahui bahwa suatu tes memiliki validitas ramalan dapat dilakukan dengan mencari korelasi antara tes hasil belajar yang sedang diuji dengan hasil belajar setelah peserta tes mengikuti pembelajaran.

(b) Validitas bandingan

Sudijono (2012:176-177) menjelaskan bahwa validitas bandingan juga dapat disebut sebagai validitas ada sekarang. Validitas ada sekarang menunjuk pada hubungan antara tes skor yang dicapai dengan keadaan sekarang. Pada keadaan ini, tes dipasangkan dengan hasil pengalaman. Pengalaman selalu

mengenai kejadian yang lampau sehingga data pengalaman tersedut sekarang sudah ada. Tes dikatakan memiliki validitas bandingan apabila tes tersebut dalam kurun waktu yang sama secara tepat telah mampu menunjukkan adanya hubungan yang searah antara tes pertama dengan tes yang dilakukan berikutnya.

b) Validitas item

Sudijono (2012: 182) mengemukakan validitas item merupakan suatu ketepatan mengukur yang dimiliki oleh instrumen tes pada tiap butirnya yang menjadi bagian tak terpisahkan dari tes secara keseluruhan, dalam mengukur apa yang seharusnya diukur oleh tes tersebut.

2. Reliabilitas

Menurut Arikunto (2013: 100) “pengertian reliabilitas tes, berhubungan dengan masalah ketepatan hasil tes”. Didukung dengan pendapat Arifin (2014: 258) bahwa reliabilitas merupakan tingkat atau derajat konsistensi dari suatu alat pengukur atau instrumen. Berdasarkan kedua pendapat tersebut, maka reliabilitas dapat diartikan sebagai tingkat keajegan hasil pengukuran dengan tes yang sama pada waktu yang berbeda. Sebuah instrumen dikatakan memiliki tingkat reliabilitas yang tinggi apabila dapat dengan ajeg memberikan data yang sesuai dengan kenyataan.

Menurut Arikunto (2013: 100-101) ajeg atau tetap tidak harus selalu sama, tetapi mengikuti perubahan secara ajeg. Ajeg disini memiliki arti sama dalam kedudukan siswa diantara anggota kelompok yang lain. Tentu saja tidak dituntut selalu tetap kedudukannya. Besar ketetapannya inilah yang menunjukkan tingkat

reliabilitas sebuah instrumen . Arikunto (2013: 101-104) memaparkan bahwa hal-hal yang mempengaruhi hasil tes dapat dikelompokkan menjadi 3 sebagai berikut:

- a) Hal yang berhubungan dengan tes itu sendiri, yaitu panjang tes dan kualitas butir-butir soalnya. Tes dengan jumlah yang banyak butirnya tentu saja lebih valid dibandingkan dengan tes yang hanya terdiri dari beberapa butir soal. Semakin panjang tes maka reliabilitasnya semakin tinggi.
- b) Hal yang berhubungan dengan terdoba (*testee*). Suatu tes yang dicobakan kepada peserta didik yang jumlahnya banyak akan mencerminkan keragaman hasil yang menggambarkan besar kecilnya reliabilitas tes.
- c) Hal yang berhubungan dengan penyelenggaraan tes. Penyelenggaraan tes yang bersifat administratif sangat menentukan hasil tes. Hasil tes tersebut akan mempengaruhi reliabilitas soal tes.

Menurut Arikunto (2013: 104) ada tiga macam metode dalam menghitung reliabilitas, yaitu:

- a) Metode Bentuk Paralel

Tes paralel merupakan dua buah tes yang memiliki kesamaan tujuan, tingkat kesukaran dan susunan yang sama, tetapi butir-butir soalnya berbeda. Kelemahan dari metode ini adalah memerlukan waktu yang cukup lama karena harus menyusun dua buah tes dan dicoba sebanyak dua kali juga.

- b) Metode Tes Ulang

Pada metode ini, tes diujikan sebanyak dua kali pada kelompok peserta didik yang sama dalam waktu yang berbeda dan hanya terdiri dari satu seri tes.

Hasil dari kedua tes tersebut dihitung korelasinya untuk mendapatkan nilai reliabilitas.

c) Belah Dua

Pada metode ini, pengetes hanya menggunakan sebuah tes dan dicobakan satu kali. Oleh karena itu disebut juga *single-test-single-trial method*. Pada waktu membelah dua dan mengkorelasikan dua belahan, baru diketahui reliabilitas separo tes.

3. Tingkat Kesukaran

Arikunto (2013: 222) menjelaskan bahwa soal yang baik adalah soal yang tidak terlalu mudah dan tidak terlalu sulit. Soal yang terlalu mudah tidak dapat merangsang siswa untuk mengembangkan kemampuannya dalam memecahkan soal tersebut. Soal yang terlalu sulit akan membuat siswa menjadi putus asa untuk mencoba lagi karena di luar kemampuan siswa.

4. Daya Pembeda

Daryanto (2008: 183) menjelaskan bahwa daya pembeda merupakan kemampuan pada setiap butir soal untuk membedakan antara siswa yang mempunyai kemampuan tinggi dan kemampuan rendah. Didukung dengan pendapat Arifin (2014: 273) yang memaparkan bahwa “perhitungan daya pembeda adalah pengukuran sejauh mana suatu butir soal mampu membedakan peserta didik yang sudah menguasai kompetensi dengan peserta didik yang belum atau kurang menguasai kompetensi berdasarkan kriteria tertentu”. Perhitungan daya pembeda dibedakan antara kelompok kecil dan kelompok besar. Kelompok

kecil merupakan kelompok yang terdiri kurang dari 100 (seratus) orang, sebaliknya kelompok besar adalah kelompok yang terdiri lebih dari 100 (seratu) orang.

a) Untuk kelompok kecil

Seluruh kelompok peserta tes (*testee*) dibagi dua sama besar, 50% kelompok atas (JA) dan kelompok bawah (JB). Seluruh pengikut tes, diderutkan mulai dari skor teratas sampai terbawah lalu dibagi dua.

b) Untuk kelompok besar

Mengingat biaya dan waktu untuk menganalisis, maka untuk kelompok besar biasanya hanya diambil kedua kutubnya saja, yaitu 27% skor teratas sebagai kelompok atas (JA) dan 27% skor terbawah sebagai kelompok bawah (JB). (Arikunto, 2013: 227).

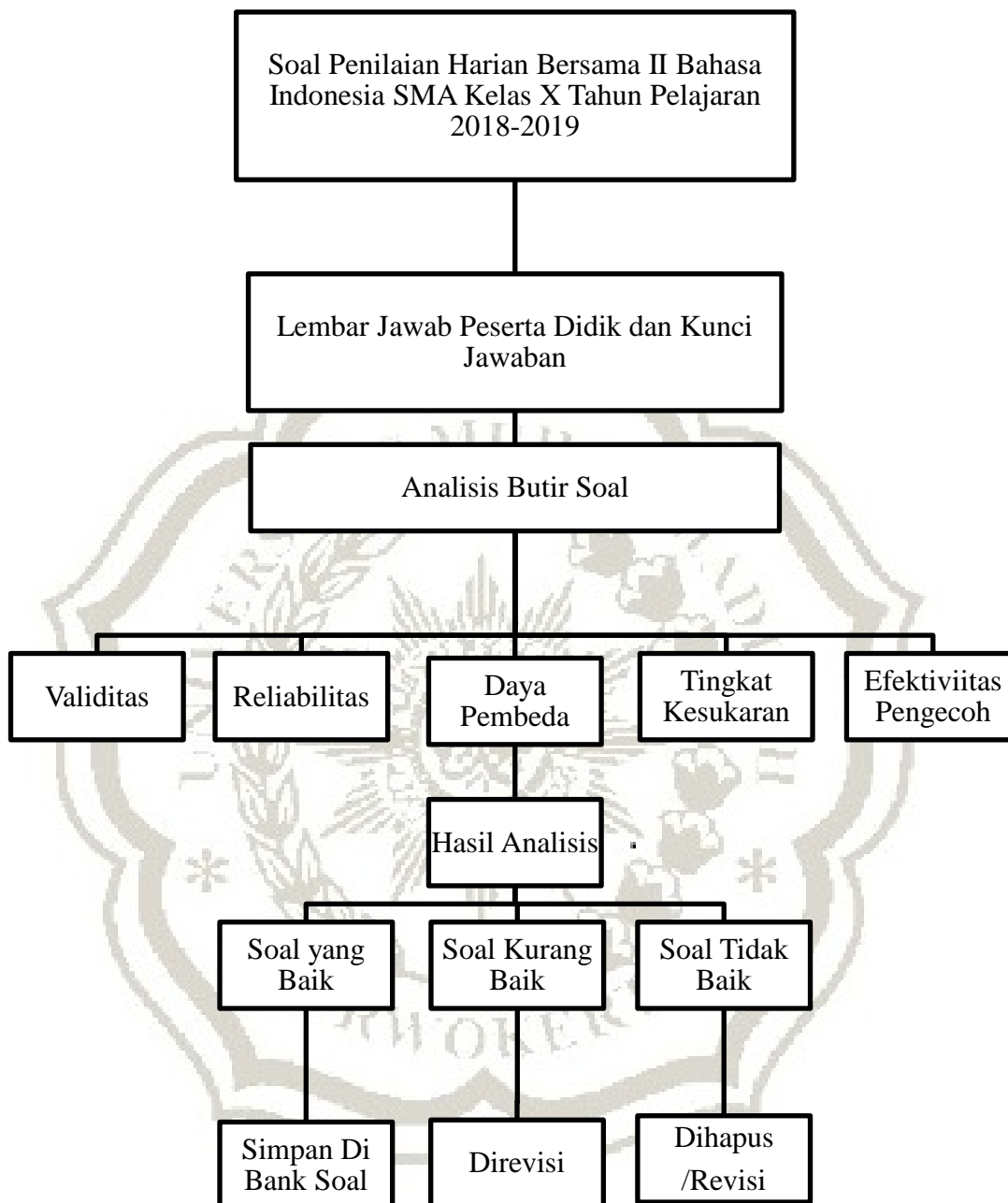
5. Efektivitas Pengecoh

Pengecoh merupakan *option* atau pilihan jawaban yang lain dari jawaban yang benar. Suatu *option* disebut efektif jika memenuhi fungsinya atau tujuan disajikannya *option* tersebut tercapai. Hal ini berarti bahwa setiap *option* yang disajikan masing-masing mempunyai kemungkinan yang sama untuk dipilih, jika tes menjawab soal itu dengan menerka-nerka (spekulasi). Pengecoh dikatakan dapat menjalankan fungsinya dengan baik apabila pengecoh tersebut dipilih sekurang-kurangnya 5% dari seluruh peserta tes. Pengecoh yang telah menjalankan fungsinya dengan baik dapat digunakan kembali pada tes yang akan datang.

C. Kerangka Pikir

Dari uraian yang telah penulis paparkan di atas maka dapat dikatakan bahwa, soal dikatakan baik (memenuhi syarat) dapat dilihat dari beberapa hal. Hal tersebut adalah validitas (kesahihan) atau sering disebut dengan apa yang akan diukur. Keterpercayaan (reliabilitas) mengacu bahwa suatu alat ukur (tes) akan menghasilkan suatu yang kurang lebih sama apabila dikenakan pada setiap individu meskipun berbeda pelaksanaannya. Indeks tingkat kesukaran mengacu bahwa suatu tes tidak terlalu sukar dan juga tidak terlalu mudah. Indeks daya beda menunjukkan bahwa sebuah butir soal harus bisa membedakan antara siswa yang berkemampuan tinggi dan siswa yang berkemampuan rendah. Efektivitas pengecoh mengisyaratkan bahwa semua pilihan jawaban (opsi) harus ada yang memilih dan jawaban betul lebih banyak dipilih oleh siswa yang berkemampuan tinggi.

Oleh karena itu ketika butir soal sudah dapat dikatakan baik (memenuhi kriteria) maka butir-butir soal itu dapat disimpan di bank soal dan dapat digunakan lagi. Sedangkan butir-butir soal yang belum dapat dikatakan baik dapat direvisi kembali dan butir-butir soal yang tidak memenuhi kriteria harus dihapus. Berikut ini merupakan bagan kerangka pikir analisis butir soal:



Gambar 1. Kerangka Pikir Analisis Butir Soal