

BAB II KAJIAN TEORITIK

A. Deskripsi Konseptual

1. Evaluasi Pembelajaran Matematika

Dalam suatu sistem pembelajaran, evaluasi merupakan salah satu komponen penting dan tahap yang perlu dilakukan guru untuk mengetahui keefektifan pembelajaran. Hasil yang diperoleh siswa dapat dijadikan sebagai umpan balik bagi guru dalam memperbaiki dan menyempurnakan program dan kegiatan pembelajaran. Guru sering memberikan ulangan harian, ulangan akhir semester, tes lisan, tes tertulis dan lain-lain, istilah-istilah tersebut pada dasarnya merupakan bagian dari sistem evaluasi (Arifin, 2012: 6). Menurut Sukardi (2008: 12), evaluasi merupakan bagian dari proses belajar mengajar yang secara keseluruhan tidak dapat dipisahkan dari kegiatan mengajar.

Berdasarkan pengertian di atas, evaluasi merupakan kegiatan yang dilakukan guru untuk mengetahui sejauhmana siswa memahami materi pembelajaran sesuai dengan tujuan. Evaluasi pembelajaran yang dilakukan oleh guru salah satunya adalah dengan mengerjakan beberapa soal. Ulangan Akhir Semester, Ulangan Harian, Ulangan Kenaikan Kelas, Ujian Nasional dan lain-lain merupakan beberapa alat evaluasi yang umum dilaksanakan di sekolah.

Alat evaluasi yang digunakan untuk penilaian mencakup tiga komponen yaitu penilaian kompetensi sikap, pengetahuan, dan

ketrampilan. Penilaian pendidikan sebagai proses pengumpulan dan pengolahan informasi untuk mengukur pencapaian hasil belajar peserta didik, memerlukan instrumen yang baik yaitu substansi, konstruksi dan penggunaan bahasa (Permendikbud, 2013).

Setiap pertanyaan hendaknya diubah menjadi pertanyaan yang lebih komunikatif yang mengukur pemahaman matematika siswa. Menurut Wardhani dan Rumiati (2011:58), secara umum pembelajaran matematika di sekolah maupun di rumah, perlu dipicu dengan soal-soal matematika yang membutuhkan penalaran, mencoba-coba (*trial*), cara penyelesaian tak tunggal, jawaban tak tunggal, dan/atau cara penyelesaian tak trivial, yang memicu keingintahuan siswa pada gagasan-gagasan matematika yang berkaitan.

2. Pentingnya Penilaian Kualitas Soal

Menurut Idowu, E.O., dkk, (2011), penilaian merupakan dasar pada pembentukan kualitas prestasi siswa di kelas maka dari itu, perlu didesain soal atau instrumen yang baik. Dengan menggunakan soal yang berkualitas baik, dapat dijadikan sebagai salah satu penilaian prestasi belajar siswa secara akurat

Mengukur kesesuaian, efisiensi dan kemantapan suatu instrumen evaluasi atau suatu tes dipergunakan bermacam-macam kualitas seperti validitas, reliabilitas, obyektivitas, dan kepraktisan. Obyektivitas adalah kualitas yang menunjukkan identitas atau kesamaan dari skor-skor atau diagnosis-diagnosis yang diperoleh dari data yang sama. Sedangkan

kepraktisan adalah suatu kualitas yang menunjukkan kemungkinan dapat dijalankannya suatu kegunaan umum dari suatu teknik penilaian dengan mendasarkan pada biaya, waktu yang diperlukan untuk menyusun, kemudahan penyusunan, mudahnya penskoran dan mudahnya interpretasian (Purwanto, 2010: 137).

Menurut Arifin (2013:69), instrumen yang baik adalah instrumen yang memenuhi syarat-syarat atau kaidah-kaidah tertentu, dapat memberikan data yang akurat sesuai dengan fungsinya, dan hanya mengukur perilaku tersebut. Sedangkan menurut Arikunto (2010), kriteria soal yang berkualitas apabila soal tersebut memenuhi persyaratan tes yaitu, validitas, reliabilitas, objektivitas, praktibilitas dan ekonomis.

Berdasarkan hal tersebut dapat disimpulkan bahwa untuk mengetahui kualitasnya suatu soal perlu dilakukan analisis secara teoritik dan analisis secara empirik. Analisis soal secara teoritik digunakan untuk mengetahui distribusi dari aspek kognitif, validitas isi, dan validitas konstruk. Namun pada dasarnya menurut Arifin (2013:132), penelaahan soal secara teoritik atau rasional yaitu dari materi, konstruk, dan bahasa dilakukan sebelum soal tersebut diujikan.

Sedangkan analisis secara empirik untuk menganalisis hasil ujian atau hasil uji-coba secara kuantitatif. Menurut Arifin (2013: 266-281), analisis kualitas butir soal dapat dilakukan dengan menghitung daya pembeda, tingkat kesukaran, homogenitas soal, dan efektivitas opsi. Arifin (2012: 149) juga mempertegas dalam bukunya bahwa, untuk mengetahui

kualitas soal uraian dapat dilakukan dengan menghitung daya pembeda dan tingkat kesukaran soal. Untuk mengetahui hal tersebut maka perlu dilakukan analisis setelah soal tersebut diujikan.

Menurut Purwanto (2010: 118), tujuan khusus dari analisis butir soal adalah untuk mengetahui soal tes mana yang baik dan mana yang belum baik, dan mengapa item atau soal itu dikatakan baik dan belum baik. Dengan mengetahui soal-soal yang belum baik itu selanjutnya kita dapat mencari kemungkinan sebab-sebab mengapa item itu belum baik. Membuat analisis soal, sedikitnya kita dapat mengetahui tiga hal penting yang dapat diperoleh dari tiap soal, yaitu:

- a. Sampai dimana tingkat atau taraf kesukaran soal itu (*difficulty level of an item*).
- b. Apakah soal itu mempunyai daya pembeda (*discriminating power*), sehingga dapat membedakan kelompok siswa yang pandai dengan kelompok siswa yang kurang pandai.
- c. Apakah semua alternatif jawaban (*options*) menarik jawaban-jawaban ataukah ada yang demikian tidak menarik sehingga tidak perlu dimasukkan ke dalam soal.

Berdasarkan uraian di atas, kualitas soal yang diukur dalam penelitian ini adalah untuk mengetahui tingkat kesukaran, daya pembeda, efektifitas distraktor dan reliabilitas. Analisis setiap butir soal tersebut bertujuan untuk mengetahui apakah soal tersebut perlu disimpan atau diperbaiki.

3. Ulangan Akhir Semester

Berdasarkan Permendikbud 2013, Ulangan Akhir Semester (UAS) adalah kegiatan yang dilakukan oleh pendidik untuk mengukur pencapaian kompetensi peserta didik di akhir semester. Cakupan ulangan meliputi seluruh indikator yang merepresentasikan semua kompetensi dasar pada semester tersebut. Pada dasarnya sekolah diberi kebijakan untuk melaksanakan UAS sendiri namun, dikoordinasikan dengan dinas pendidikan, kementerian agama, atau organisasi seperti: K3S, MKKS, KKMI, atau KKKMI.

Soal Ulangan Akhir Semester gasal matematika di kabupaten Banjarnegara tahun pelajaran 2014/2015 terdiri dari tes obyektif dan uraian. Tes obyektif tersebut menggunakan tipe soal pilihan ganda.

a. Pilihan Ganda

Menurut Arifin (2012:156), soal tes bentuk pilihan-ganda dapat digunakan untuk mengukur hasil belajar yang lebih kompleks dan terdiri atas pembawa pokok persoalan dan pilihan jawaban. Sedangkan menurut Arikunto (2010:168), tes pilihan ganda terdiri atas suatu pengertian yang belum lengkap sehingga harus memilih satu jawaban yang telah disediakan.

Berdasarkan pengertian tersebut soal tes pilihan ganda merupakan tes yang terdiri dari pertanyaan yang terdapat beberapa pilihan jawaban dan hanya terdapat satu kunci jawaban. Pengecoh atau distraktor yang digunakan hendaknya dapat berfungsi dengan baik atau

memiliki daya tarik bagi peserta tes yang kurang memahami materi untuk memilih alternatif jawaban tersebut. Pada prinsipnya soal pilihan ganda terdiri atas sebuah pokok persoalan atau problem dengan alternatif pilihan jawaban yang dianjurkan diisi oleh siswa untuk dievaluasi. Selain itu setiap butir soal dibedakan dalam dua bagian penting, yaitu pokok persoalan dan jawaban alternatif.

b. Uraian

Tes uraian merupakan pertanyaan yang menuntut siswa untuk menjawabnya dalam bentuk menjelaskan, menguraikan, membandingkan, memberi alasan, dan bentuk lain yang sejenis dengan tuntutan soal. Menurut Rasyid, Harun dan Mansur (2009:190), tes uraian objektif sangat tepat digunakan dalam bidang matematika dan sains karena memiliki kunci jawaban hanya satu. Pengerjaan soal ini menggunakan prosedur atau langkah-langkah tertentu. Soal bentuk uraian, memungkinkan siswa menjawab menggunakan bahasa mereka sendiri namun, harus ada pokok-pokok penting dalam sistematika jawabannya.

4. Domain Kognitif

Rumusan domain kognitif yang dikemukakan oleh Mullis, dkk., (2009: 20), mengklasifikasikan bahwa terdapat tiga tingkatan kognitif siswa dalam belajar matematika yaitu pengetahuan (*knowing*), penerapan (*applying*), penalaran (*reasoning*). Tabel berikut menunjukkan kemampuan pada dimensi kognitif (Mullis, dkk., 2009: 42-46) :

Domain Kognitif	Indikator Kriteria
Pemahaman (<i>knowing</i>)	<ol style="list-style-type: none"> (1). Mengingat, yaitu memahami definisi, sifat-sifat peristilahan, notasi-notasi dalam matematika. (2). Mengenali, yaitu mengenal objek dan entitas matematika, (3). Menghitung, yaitu menghitung prosedur-prosedur algoritmik, (4). Mengukur, yaitu menggunakan instrument-instrumen pengukuran, (5). Mengklasifikasi, yaitu mengklasifikasikan objek, bangun, bilangan, serta mengurutkan objek, bangun, dan bilangan berdasarkan sifat-sifat tertentu. (6). Mengurutkan, yaitu mengambil informasi dari grafik, tabel, atau sumber lain yang sederhana.
Penerapan (<i>applying</i>)	<ol style="list-style-type: none"> (1). Memilih, yaitu memilih operasi, metode, serta strategi yang tepat untuk memecahkan masalah dimana prosedur, algoritma, atau metode untuk menyelesaikan masalah tersebut sudah diketahui, (2). Merepresentasi, yaitu menyajikan informasi matematika atau data dalam bentuk diagram, tabel, dan grafik, (3). Memodelkan, yaitu menggunakan model seperti persamaan, bidang geometris, atau diagram untuk memecahkan masalah rutin, (4). Mengimplementasi yaitu mengimplementasikan sekumpulan perintah-perintah matematika, seperti menspesifikasikan bentuk dan diagram, (5). Memecahkan masalah rutin, yaitu menyelesaikan masalah matematika dalam konteks yang sederhana atau familiar.
Penalaran (<i>reasoning</i>)	<ol style="list-style-type: none"> (1). Menganalisa, yaitu menentukan, mendiskripsikan, atau menggunakan hubungan antar variabel atau objek dalam situasi matematika, serta membuat kesimpulan yang tepat terhadap informasi yang diberikan, (2). Menggeneralisasi yaitu memperluas ruang lingkup dari hasil pemikiran matematika dan menegaskan kembali hasilnya dalam istilah yang lebih umum, (3). Mensintesis, yaitu membuat hubungan dari elemen-elemen yang berbeda serta

	<p>menghubungkan ide-ide matematika yang saling berdekatan. Selain itu, pada tingkatan ini juga menggabungkan fakta matematika, konsep, dan prosedur untuk menetapkan hasil, dan menggabungkannya untuk mendapatkan hasil lebih lanjut,</p> <p>(4). Memberi alasan yaitu membuat pembenaran atau membuktikan dengan menggunakan hasil maupun sifat-sifat matematika yang telah diketahui,</p> <p>(5). Menyelesaikan masalah non rutin yaitu menyelesaikan sekumpulan masalah matematika atau masalah kehidupan sehari-hari yang kompleks. Selain itu, pada tingkatan ini siswa mampu mengaplikasikan fakta, prosedur, dan konsep pada permasalahan yang kompleks atau tidak familiar.</p>
--	---

Tabel. 2.1 : Kemampuan Domain Kognitif

Soal-soal matematika diharapkan mampu mengukur tingkat kemampuan siswa dari mengetahui fakta, prosedur atau konsep hingga menggunakannya untuk memecahkan masalah yang sederhana sampai masalah yang memerlukan penalaran tinggi. Domain kognitif yang dirumuskan oleh Mullis, dkk., (2009) tidak jauh berbeda dengan tingkatan kognitif yang rumuskan oleh Anderson dan Krathwohl (2001) yang merevisi domain kognitif dari Bloom. Namun, menurut Anderson dan Krathwohl (2001), tingkatan kognitif dibagi menjadi enam kategori, yaitu mengingat (*remember*), memahami (*understand*), mengaplikasikan (*applying*), menganalisis (*analyze*), mengevaluasi (*evaluation*), dan membuat (*create*).

Penelitian ini menggunakan domain kognitif yang dirumuskan oleh Mullis, dkk., (2009) karena pada kriteria indikator masing-masing domain, lebih terfokus pada pembelajaran matematika kelas VIII. Hal tersebut lebih mendukung soal yang hendak dianalisis oleh peneliti.

5. Analisis Butir Soal

Analisis butir soal atau analisis item merupakan penyajian pertanyaan-pertanyaan tes agar diperoleh perangkat pertanyaan yang memiliki kualitas memadai (Sudjana, 2010: 135). Dengan menggunakan analisis butir soal, maka dapat diketahui kualitas soalnya, sehingga dapat dijadikan sebagai evaluasi dalam penyusunan soal. Analisis butir soal dijabarkan sebagai berikut:

a. Tingkat Kesukaran

Tingkat kesukaran soal dilihat dari kesanggupan atau kemampuan siswa dalam menjawab, bukan dilihat dari sudut pandang guru sebagai pembuat soal. Hal penting yang perlu diperhatikan dalam melakukan analisis tingkat kesukaran soal adalah penentuan proporsi dan kriteria soal yang termasuk mudah, sedang dan sukar (Sudjana, 2010: 135).

Pertimbangan menentukan jumlah soal kategori mudah, sedang dan sukar, terdapat kriteria yang perlu diperhatikan. Pertimbangan pertama adalah adanya keseimbangan, yaitu jumlah sama untuk ketiga kategori tersebut. Pertimbangan yang kedua adalah, soal mudah-sedang-sukar bisa dibuat perbandingan 3-4-3. Artinya, 30% soal kategori mudah, 40% soal kategori sedang dan 30% soal kategori sukar. Pertimbangan ketiga adalah soal bisa dibuat dengan proporsi 3-5-2. Artinya, 30% soal kategori mudah, 50% soal kategori sedang, dan 20% soal kategori sukar (Sudjana, 2010: 135-136).

Perhitungan tingkat kesukaran soal adalah pengukuran seberapa besar derajat kesukaran suatu soal. Jika suatu soal memiliki tingkat kesukaran seimbang (proporsional), maka dapat dikatakan bahwa soal tersebut baik. Suatu soal tes hendaknya tidak terlalu sukar dan tidak pula terlalu mudah (Arifin, 2012: 342). Begitu pula yang diungkapkan oleh Arikunto (2010:207), soal yang baik adalah soal yang tidak terlalu mudah atau tidak terlalu sukar.

Kriteria yang digunakan adalah semakin kecil indeks yang diperoleh, maka makin sulit soal tersebut. Sebaliknya, semakin besar indeks yang diperoleh, makin mudah soal tersebut. Kriteria tingkat kesulitan soal tersebut adalah sebagai berikut:

Tingkat Kesulitan	Kriteria
0 – 0,30	soal kategori sukar
0,31 – 0,70	soal kategori sedang
0,71 – 1,00	soal kategori mudah

Tabel 2.2 : Kriteria Tingkat Kesulitan Soal

(Arifin, 2013: 273)

b. Daya Beda (*Discriminating Power*)

Analisis daya beda mengkaji butir-butir soal dengan tujuan untuk mengetahui kesanggupan soal dalam membedakan siswa yang tergolong mampu dengan siswa yang tergolong kurang atau berprestasi rendah. Artinya, bila soal tersebut diberikan kepada anak yang mampu menunjukkan hasil prestasi yang tinggi, dan bila diberikan kepada siswa yang lemah hasilnya rendah. Tes dikatakan tidak memiliki daya pembeda apabila tes tersebut jika diujikan kepada anak berprestasi tinggi hasilnya rendah tetapi, bila diberikan kepada anak yang lemah hasilnya tinggi. Atau bila

diberikan kepada keduanya, hasilnya sama. Tes yang tidak memiliki daya pembeda, tidak akan menghasilkan gambaran hasil yang sesuai dengan kemampuan siswa yang sebenarnya (Sudjana, 2010: 141).

Perhitungan daya beda adalah pengukuran sejauhmana suatu butir soal mampu membedakan peserta didik yang sudah menguasai kompetensi dengan peserta didik yang belum/kurang menguasai kompetensi berdasarkan kriteria tertentu. Semakin tinggi koefisien daya beda suatu butir soal, semakin mampu butir soal tersebut membedakan antara peserta didik yang menguasai kompetensi dengan peserta didik yang kurang menguasai kompetensi (Arifin, 2012: 350).

Daya beda suatu tes merupakan kemampuan suatu tes dapat digunakan untuk membedakan siswa yang berkemampuan rendah dengan siswa berkemampuan tinggi. Suatu soal dikatakan memiliki daya pembeda yang baik, apabila soal tersebut jika dikerjakan oleh siswa berkemampuan tinggi sebagian besar menjawab benar dan jika dikerjakan oleh siswa berkemampuan rendah sebagian besar menjawab salah.

Daya Pembeda	Kriteria
0.40 ke atas	Sangat baik
0.30 – 0.39	Baik
0.20 – 0.29	Cukup, soal perlu perbaikan
0,19 ke bawah	Kurang baik

Tabel. 2.3 : Kriteria Daya Beda Soal

(Arifin, 2013: 274)

c. Analisis Pengecoh (Distraktor)

Distraktor merupakan pengecoh dari suatu jawaban yang berbentuk objektif. Pola jawaban soal diperoleh dengan menghitung banyaknya responden yang memilih jawaban a, b, c atau d atau yang tidak memilih pilihan manapun (blangko).

Menurut Arifin (2012: 357), pengecoh dianggap baik bila jumlah peserta didik yang memilih pengecoh itu sama atau mendekati jumlah ideal, yaitu 5% dari jumlah peserta tes. Pada soal bentuk pilihan ganda ada alternatif jawaban (opsi) yang merupakan pengecoh. Butir soal yang baik, pengecohnya akan dipilih secara merata oleh peserta didik yang menjawab salah. Sebaliknya, butir soal yang kurang, pengecohnya akan dipilih secara tidak merata.

d. Reliabilitas

Menurut Sudjana (2010: 148), analisis reliabilitas suatu tes dan atau alat ukurnya, merupakan nontes, pada hakikatnya menguji keajegan pertanyaan tes apabila diberikan berulang kali kepada objek yang sama. Analisis reliabilitas suatu tes dan atau alat ukur lainnya pada umumnya menggunakan teknik korelasi.

Sedangkan menurut Arifin (2012: 326), Reliabilitas suatu tes adalah tingkat atau derajat konsistensi tes yang bersangkutan. Reliabilitas berkenaan dengan pertanyaan, apakah suatu tes teliti dan dapat dipercaya sesuai dengan kriteria yang telah ditetapkan. Suatu tes dapat dikatakan reliabel jika selalu memberikan hasil

yang sama bila diteskan pada kelompok yang sama pada waktu atau kesempatan yang berbeda.

Reliabilitas adalah keajegan. Suatu tes dikatakan reliabel apabila dilakukan beberapa kali pengujian menunjukkan hasil yang relatif sama. Pengujian suatu tes dapat dilakukan terhadap objek yang sama pada waktu berbeda dengan selang waktu yang tidak terlalu lama dan juga terlalu singkat, bisa juga dilakukan dengan membandingkan hasil pengujian dari tes yang setara. Kriteria reliabilitas soal dapat diklasifikasikan sebagai berikut:

Reliabilitas	Kriteria
$0.80 < r \leq 1.00$	Sangat tinggi
$0.60 < r \leq 0.80$	Tinggi
$0.40 < r \leq 0.60$	Cukup
$0.20 < r \leq 0.40$	Rendah
$0.000 < r \leq 0.20$	Sangat rendah

Tabel 2.4 : Kriteria Reliabilitas Soal

(Priatna, 2008: 16)

e. *Item and Analysis Test (Iteman)*

Menurut Kartowagiran (2011), Iteman merupakan analisis butir secara empirik dengan pendekatan teori tes klasik. Sedangkan menurut Hidayati (2010), Iteman merupakan perangkat lunak (*software*) yang dibuat melalui bahasa pemrograman komputer dan dibuat khusus untuk menganalisis butir soal dan tes. Berdasarkan pengertian di atas maka Iteman adalah salah satu aplikasi komputer yang digunakan untuk menganalisis butir soal dan tes. Hasil analisis Iteman ini meliputi: tingkat kesukaran butir soal, daya

pembeda soal, statistik sebaran jawaban, reliabilitas tes, kesalahan pengukuran (*standar error*), dan distribusi skor setiap peserta tes.

Hasil dari analisis Iteman ini berupa dua file yaitu, file statistik dan file skor. Keduanya merupakan file ASCII yang dapat dilihat dengan menggunakan program pengolahan kata (*word procesor*). File skor memberikan hasil skor untuk setiap peserta tes yang menunjukkan jumlah benar dari seluruh jawaban. File statistik hasil analisis dari Iteman dapat dibedakan ke dalam dua bagian, yaitu: statistik butir soal dan statistik tes (skala).

(a) Statistik Butir Soal

Statistik butir-butir soal yang bersifat dikotomi yaitu pilihan ganda, statistik berikut adalah *output* dari setiap butir soal yang dianalisis:

- (1) *Seq. No* adalah nomor urut butir soal dalam file data.
- (2) *Scala-item* adalah nomor urut butir soal dalam skala (tes/subtes).
- (3) *Prop. Correct* adalah proporsi siswa (peserta tes) yang menjawab benar butir soal.
- (4) *Biser* adalah indeks daya pembeda soal dengan menggunakan koefisien korelasi biserial.
- (5) *Point-biser* adalah juga indeks daya pembeda soal dan pilihan jawaban (alternatif) dengan menggunakan koefisien korelasi point-biserial.

(6) Statistik pilihan jawaban (alternatif) memberikan informasi yang sama dengan statistik butir soal. Tanda bintang (*) yang muncul di sebelah kanan hasil analisis menunjukkan kunci jawaban.

(b) Statistik Tes/Skala

Hasil analisis statistik tes/skala adalah sebagai berikut:

- (1) *N of items* adalah jumlah butir soal dalam tes/skala yang ikut dianalisis.
- (2) *N of examines* adalah jumlah peserta tes yang digunakan dalam analisis.
- (3) *Mean* adalah skor rata-rata peserta tes.
- (4) *Variance* adalah varian dari distribusi skor peserta tes yang memberikan gambaran tentang sebaran skor peserta tes.
- (5) *Std. Dev* adalah deviasi standar dari distribusi skor peserta tes. Deviasi standar adalah akar dari variance.
- (6) *Alpha* adalah koefisien reliabilitas alpha untuk tes/skala tersebut yang merupakan indeks homogenitas tes/skala.
- (7) *Minimum* adalah skor terendah peserta tes dalam tes/skala tersebut.
- (8) *Maximum* adalah skor tertinggi peserta tes dalam tes/skala tersebut.

B. Penelitian yang Relevan

Beberapa peneliti menunjukkan pentingnya menganalisis distribusi domain kognitif dan kualitas soal dalam pembelajaran matematika bagi guru dan siswa. Rahayu, dkk., (2012) dalam kajiannya terhadap soal-soal geometri menunjukkan bahwa aspek kognitif (*required behavior*) yang termasuk dalam domain *reasoning* sangat sedikit dibandingkan dengan aspek kognitif (*required behavior*) pada domain *knowing* dan *applying* yakni 5.07%.

Hasil penelitian yang dilakukan oleh Aningtyas (2012), dengan secara teoritik untuk mengetahui aspek kognitif yang terdapat pada soal Ulangan Akhir Semester matematika tahun ajaran 2010/2011 dan 2011/2012 tidak tersebar merata, yaitu hanya aspek mengingat (*remembering*), memahami (*understanding*), mengaplikasikan (*applying*) dan menganalisis (*analyzing*). Aspek mengaplikasikan (*applying*) merupakan aspek kognitif yang paling banyak dijumpai pada soal-soal yaitu pada soal Ulangan Akhir Semester Gasal Tahun Ajaran 2010/2011 sebesar 64,44% dan soal Ulangan Akhir Semester Gasal Tahun Ajaran 2011/2012 sebesar 62,22%. Sedangkan aspek mengevaluasi (*evaluating*) dan mengkreasi (*creating*) tidak terdapat pada soal ulangan akhir semester tersebut.

Sementara itu penelitian yang dilakukan oleh Wardaningsih (2012), mengkaji soal Ujian Nasional matematika SMP untuk mendapatkan informasi tentang klasifikasi soal berdasarkan distribusi materi dan tingkatan kognitif, dan mengukur karakteristik soal. Untuk mengukur karakteristik soal dimulai

dengan mengumpulkan soal beserta kunci jawaban dan pembahasan, menyusun tabel perhitungan, mengujicobakan soal, melakukan analisis data, dan menarik kesimpulan.

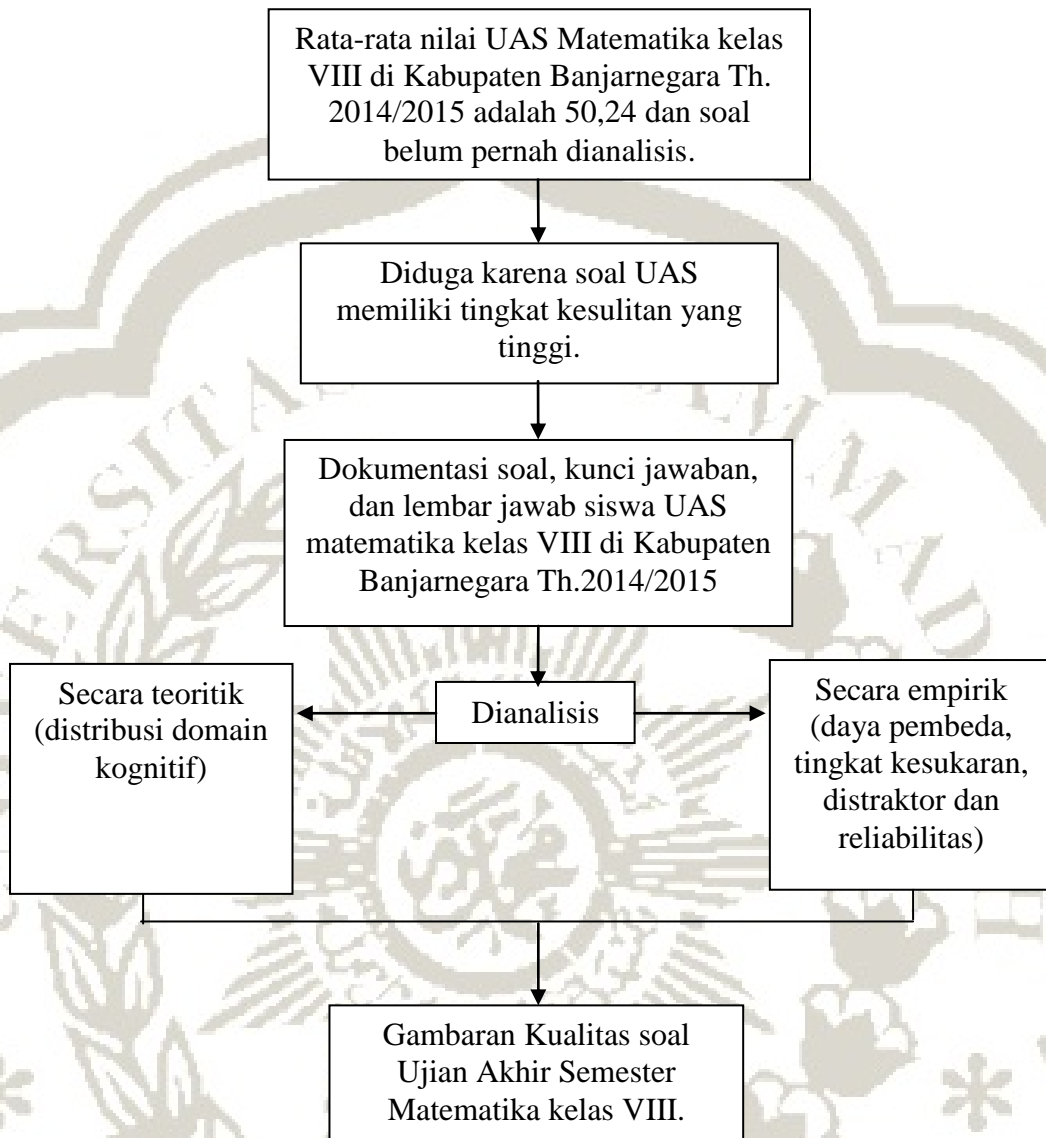
Berdasarkan tiga kajian penelitian di atas, peneliti merasa tertarik untuk melakukan penelitian tentang analisis kualitas soal berdasarkan domain kognitif dan penghitungan statistik menggunakan komputer. Pada kesempatan kali ini, peneliti akan menggunakan perhitungan secara kuantitatif untuk mendapatkan informasi tentang tingkat kesukaran, daya pembeda, efektifitas distraktor, dan reliabilitas, serta analisis secara teoritik berdasarkan domain kognitif.

C. Kerangka Berfikir

Evaluasi proses pembelajaran merupakan cara yang dilakukan agar mengetahui hasil belajar siswa sehingga dapat ditentukan tindak lanjut yang akan dilakukan. Alat evaluasi yang biasa digunakan berupa tes. Tes tersebut biasanya digunakan untuk mengukur prestasi siswa secara kognitif. Berdasarkan hasil TIMSS, menunjukkan bahwa Indonesia berada pada tingkatan kognitif mengetahui (*knowing*), yang merupakan tingkatan terendah berdasarkan domain kognitif menurut Mullis, dkk (2009). Soal seharusnya dibuat sebaik mungkin, dengan melakukan analisis tingkat kesukaran, daya pembeda, efektifitas distraktor, dan reliabilitas diharapkan dapat digunakan sebagai evaluasi dalam penyusunan soal yang baik.

Soal UAS matematika disusun oleh guru-guru berkompeten yang tergabung dalam MGMP Kabupaten Banjarnegara, soal tersebut diharapkan menghasilkan soal yang berkualitas. Berdasarkan data yang diperoleh peneliti menunjukkan bahwa rata-rata nilai UAS matematika kelas VIII dari 8 sekolah di Kabupaten Banjarnegara adalah 50,24. Hal tersebut menunjukkan bahwa 8 dari 88 siswa kelas VIII mendapatkan nilai di bawah KKM pada UAS matematika di Kabupaten Banjarnegara. Hal tersebut diduga karena soal UAS memiliki tingkat kesulitan yang tinggi.

Berdasarkan penjelasan di atas, perlu dilakukan analisis soal UAS matematika. Analisis berdasarkan domain kognitif diperlukan untuk mengetahui sebaran tingkatan kognitif yang terdapat pada soal UAS matematika di Kabupaten Banjarnegara. Dokumen soal dan lembar jawab kemudian dianalisis sehingga dapat memberikan gambaran kualitas dan level kognitif soal UAS matematika. Berdasarkan uraian di atas maka dapat disusun kerangka berfikir seperti pada diagram alur penelitian berikut:



Gambar. 2.1 : Diagram Alur Penelitian