

BAB I

PENDAHULUAN

A. Latar Belakang Masalah

Evaluasi merupakan salah satu bagian dari proses pembelajaran. Evaluasi itu sendiri adalah suatu proses yang sistematis dan berkelanjutan untuk menentukan kualitas (nilai dan arti) dari sesuatu, berdasarkan pertimbangan dan kriteria tertentu dalam rangka pembuatan keputusan (Arifin, 2009). Evaluasi dapat digolongkan menjadi dua berdasarkan waktu pelaksanaannya, yaitu evaluasi formatif dan evaluasi sumatif. Evaluasi formatif dilaksanakan pada setiap kali satuan program pelajaran atau subpokok bahasan dapat diselesaikan, sedangkan evaluasi sumatif dilaksanakan setelah sekumpulan program pelajaran atau seluruh unit pelajaran selesai diajarkan. Tujuan utama dari evaluasi sumatif ini adalah untuk menentukan nilai yang melambangkan keberhasilan peserta didik setelah mereka menempuh proses pembelajaran dalam jangka waktu tertentu, sehingga dapat ditentukan dapat tidaknya peserta didik untuk mengikuti program pengajaran berikutnya atau tingkat yang lebih tinggi (Sudijono, 2006).

Untuk melakukan suatu evaluasi sumatif, diperlukan adanya alat ukur (instrumen) baik yang berbentuk tes maupun non-tes dengan tujuan agar tidak salah dalam mengambil keputusan. Instrumen yang baik adalah instrumen yang memenuhi syarat-syarat atau kaidah-kaidah tertentu, dapat memberikan data yang akurat sesuai dengan fungsinya dan hanya mengukur sampel perilaku tertentu. Adapun karakteristik instrumen evaluasi yang baik adalah valid, reliabel, relevan,

representatif, praktis, deskriminatif, spesifik dan proporsional. Analisis kualitas tes merupakan suatu tahap yang harus ditempuh, baik tes secara keseluruhan maupun butir soal yang menjadi bagian dari tes tersebut (Arifin, 2009). Analisis soal bertujuan untuk mengadakan identifikasi soal-soal yang baik, kurang baik dan soal yang jelek. Hasil dari analisis soal ini adalah diperolehnya informasi tentang kejelekan sebuah soal dan petunjuk untuk mengadakan perbaikan (Arikunto, 2010).

Kualitas tes dapat diungkap melalui analisis butir soal secara teoretis (telaah) dan analisis empiris. Analisis butir soal secara teoretis (telaah) dilakukan untuk menilai butir soal ditinjau dari aspek materi, konstruksi, dan bahasa. Analisis empiris dapat menggunakan pendekatan tes klasik (*classical test theory*) maupun pendekatan tes modern (*item response theory*) (Nurung, 2008). Teori pengukuran klasik memiliki keterbatasan karena bersifat *group dependent* dan *item dependent* (Hambleton *et al.*, 1991). *Group dependent* artinya hasil pengukuran tergantung dari kelompok peserta yang mengerjakan tes. Jika tes diujikan kepada kelompok peserta dengan kemampuan tinggi, tingkat kesulitan butir soal akan rendah. Sebaliknya jika tes diujikan kepada peserta yang kemampuan rendah, tingkat kesulitan butir soal akan tinggi. *Item dependent* artinya hasil pengukuran tergantung dari tes mana yang diujikan. Jika tes yang diujikan mempunyai tingkat kesulitan yang tinggi, estimasi kemampuan peserta akan rendah. Sebaliknya jika tes yang diujikan tingkat kesulitannya rendah maka estimasi kemampuan peserta akan tinggi. Kelemahan pengukuran semacam ini tidak terdapat pada teori respon butir (*item response theory*).

Teori respon butir (*item response theory*) merupakan salah satu cara untuk menilai kelayakan butir dengan membandingkan rerata penampilan butir terhadap tampilan bukti kemampuan kelompok yang diramalkan oleh model. Tujuan utama teori respon butir dikembangkan adalah untuk mengatasi kelemahan teori tes klasik yang tidak *independent* terhadap kelompok peserta yang mengerjakan tes maupun terhadap tes yang diujikan (Hambleton *et al.*, 1991).

Analisis soal pada penelitian sebelumnya, yang dilakukan oleh Yulia Anggraeni hanya menggunakan teori tes klasik (*classical test theory*) dan analisis soal pada penelitian terbaru yang dilakukan oleh Ana Rofiati, sudah menggunakan teori respon butir (*item response theory*), tidak hanya menggunakan teori tes klasik (*classical test theory*), namun pada penelitian yang dilakukan oleh Ana Rofiati ini hanya dilakukan satu kali uji coba padahal jika uji coba dilakukan lebih dari satu kali hal ini membuka peluang agar soal (instrumen) dapat direvisi menjadi lebih baik dan penelitian ini juga hanya dilakukan pada kelas X dengan tujuan agar dapat menentukan program penjurusan dengan benar padahal diperlukan juga penelitian dengan menggunakan teori respon butir (*item response theory*) dan teori tes klasik (*classical test theory*) pada kelas XI dengan tujuan agar diperoleh instrumen yang baik sehingga tidak salah dalam mengambil keputusan mengenai dapat tidaknya peserta didik mengikuti program pengajaran ke tingkat yang lebih tinggi.

Berdasarkan pemaparan di atas, maka perlu dilakukan penelitian dengan judul “*Pengembangan Soal Ulangan Kenaikan Kelas SMA Kelas XI Berdasarkan Classical Test Theory dan Item Response Theory*”.

B. Rumusan Masalah

Berdasarkan latar belakang yang telah dikemukakan, masalah pokok yang hendak diungkapkan dalam penelitian ini adalah sebagai berikut :

“Apakah soal ulangan kenaikan kelas SMA Kelas XI yang dikembangkan telah memenuhi kriteria sebagai instrumen yang baik berdasarkan *classical test theory* dan *item response theory*?”

Rumusan masalah di atas dapat dijabarkan sebagai berikut :

1. Apakah soal ulangan kenaikan kelas SMA kelas XI yang dikembangkan memiliki validitas isi sebagai butir soal yang baik?
2. Apakah soal ulangan kenaikan kelas SMA kelas XI yang dikembangkan memiliki validitas, reliabilitas, daya pembeda, tingkat kesukaran dan kualitas pengecoh sebagai butir soal yang baik menurut *classical test theory*?
3. Apakah soal ulangan kenaikan kelas SMA kelas XI yang dikembangkan memiliki validitas, reliabilitas, daya pembeda (parameter a), tingkat kesukaran (parameter b) dan faktor tebakan (parameter c) sebagai butir soal yang baik menurut *item response theory*?

C. Pembatasan Masalah

Untuk mengatasi meluasnya permasalahan, maka dibuat pembatasan masalah untuk penelitian ini, yaitu:

1. Materi yang menjadi bahan soal ulangan kenaikan kelas yang dikembangkan adalah materi yang dipelajari di SMA kelas XI pada semester genap yaitu asam basa, larutan penyangga, hidrolisis garam, kelarutan dan hasil kali kelarutan serta sistem koloid.

2. Produk evaluasi yang dikembangkan berbentuk tes tertulis tipe pilihan ganda.
3. Subjek uji coba pada penelitian ini adalah siswa SMA kelas XI di salah satu SMA kluster I di Bandung sebagai peserta tes untuk uji coba terbatas dan siswa SMA kelas XII di salah satu SMA kluster I di Bandung sebagai peserta tes untuk uji lapangan.

D. Tujuan Penelitian

Penelitian ini bertujuan untuk menghasilkan produk berupa soal ulangan kenaikan kelas SMA kelas XI yang memiliki validitas, reliabilitas, daya pembeda, tingkat kesukaran dan kualitas pengecoh berdasarkan *classical test theory* serta validitas, reliabilitas, daya pembeda (parameter a), tingkat kesukaran (parameter b) dan faktor tebakan (parameter c) berdasarkan *item response theory* sebagai butir soal yang baik.

E. Manfaat Penelitian

Hasil penelitian ini diharapkan dapat digunakan sebagai bahan pertimbangan, antara lain:

1. Bagi peserta didik

Instrumen yang dikembangkan dapat digunakan untuk mengetahui kedudukan peserta didik dibandingkan dengan peserta didik yang lainnya dan untuk mengetahui dapat tidaknya peserta didik mengikuti program pengajaran ke tingkat yang lebih tinggi.

2. Bagi guru

Instrumen yang dikembangkan dapat digunakan untuk perbaikan dan pengembangan sistem pengajaran, mengetahui tingkat penguasaan peserta didik terhadap suatu materi dan mengurangi kesalahan dalam mengambil keputusan.

3. Bagi peneliti

Untuk menambah wawasan mengenai pengembangan tes berdasarkan *classical test theory* dan *item response theory*.

F. Definisi Operasional

Untuk menghindari perbedaan persepsi dari kajian yang dilakukan, maka peneliti perlu untuk menjelaskan istilah-istilah sebagai berikut :

1. Pengembangan tes

Pengembangan tes adalah suatu proses perancangan dan perbaikan alat ukur (tes) agar menjadi suatu alat ukur (tes) yang berkualitas (Firman, 2000).

2. Analisis soal

Analisis soal bertujuan untuk mengadakan identifikasi soal-soal yang baik, kurang baik dan soal yang jelek. Dengan analisis soal dapat diperoleh informasi tentang kejelekan sebuah soal dan petunjuk untuk mengadakan perbaikan (Arikunto, 2010).

3. Ulangan kenaikan kelas

Menurut Permendiknas No.20 tahun 2007 tentang Standar Penilaian Pendidikan, ulangan kenaikan kelas adalah kegiatan yang dilakukan oleh pendidik di akhir semester genap untuk mengukur pencapaian kompetensi peserta didik di akhir semester genap pada satuan pendidikan yang menggunakan sistem paket.

Cakupan ulangan meliputi seluruh indikator yang merepresentasikan KD pada semester tersebut.

4. *Classical test theory* (CTT) atau teori tes klasik

Classical test theory (CTT) atau teori tes klasik menurut Hambleton *et al.* (1991) adalah teori tentang skor tes yang memperkenalkan tiga konsep yaitu skor tes atau skor amatan (*test score/observed score*), skor murni (*true score*) dan skor kesalahan pengukuran (*error score*).

5. *Item response theory* (IRT) atau teori respon butir

Item Response Theory (IRT) atau teori respon butir adalah teori yang menyatakan tentang suatu pembuatan perangkat ukur yang analisisnya berdasarkan bentuk lengkungan kurva karakteristik butir yang diperoleh dari hasil respon (jawaban) para peserta tes (Susetyo, 2011).