

# BAB I

## PENDAHULUAN

### A. Latar Belakang

Pendidikan adalah suatu usaha yang dilakukan untuk mengembangkan kemampuan dan kepribadian individu melalui proses atau kegiatan tertentu (pengajaran, bimbingan atau latihan) serta interaksi individu dengan lingkungannya untuk mencapai manusia seutuhnya (Arifin, 2009). Menurut Permendiknas No.41 Tahun 2007 tentang Standar Proses, salah satu prinsip penyelenggaraan pendidikan untuk dijadikan landasan reformasi pendidikan adalah pendidikan diselenggarakan sebagai proses pembudayaan dan pemberdayaan peserta didik yang berlangsung sepanjang hayat. Implikasi dari prinsip ini adalah pergeseran paradigma proses pendidikan, yaitu dari paradigma pengajaran ke paradigma pembelajaran. Pembelajaran adalah proses interaksi peserta didik dengan guru dan sumber belajar pada suatu lingkungan belajar. Proses pembelajaran perlu direncanakan, dilaksanakan, dinilai, dan diawasi agar terlaksana secara efektif dan efisien.

Dalam sistem pembelajaran, evaluasi merupakan salah satu komponen penting dan tahap yang harus ditempuh oleh guru untuk mengetahui keefektifan pembelajaran. Hasil yang diperoleh dari evaluasi dapat dijadikan umpan balik (*feed-back*) bagi guru dalam memperbaiki dan menyempurnakan program dan kegiatan pembelajaran (Arifin, 2009). Undang-undang Nomor 20 Tahun 2003 Pasal 57 (ayat 2) menyatakan bahwa evaluasi (penilaian) dilakukan terhadap

peserta didik, lembaga dan program pendidikan jalur formal dan nonformal untuk semua jenjang, satuan dan jenis pendidikan.

Menurut Peraturan Pemerintah Nomor 19 Tahun 2005 Tentang Standar Nasional Pendidikan, penilaian adalah proses pengumpulan dan pengolahan informasi untuk mengukur pencapaian hasil belajar peserta didik. Penilaian hasil belajar oleh pendidik dilakukan secara berkesinambungan untuk memantau proses, kemajuan, dan perbaikan hasil dalam bentuk ulangan harian, ulangan tengah semester, ulangan akhir semester dan ulangan kenaikan kelas. Penilaian digunakan untuk menilai pencapaian kompetensi peserta didik, bahan penyusunan laporan kemajuan hasil belajar dan memperbaiki proses pembelajaran.

Hasil evaluasi dapat digunakan untuk menyeleksi, baik ketika peserta didik akan masuk sekolah/jenjang atau jenis pendidikan tertentu, selama mengikuti program pendidikan, pada saat akan menyelesaikan jenjang pendidikan, maupun ketika masuk dunia kerja (Arifin, 2009). Ulangan kenaikan kelas sebagai bentuk evaluasi terutama untuk tingkat SMA kelas X memiliki fungsi penting dalam proses seleksi peserta didik untuk penentuan program penjurusan yang terdiri dari Ilmu Alam, Ilmu Sosial atau Bahasa.

Dalam proses evaluasi pembelajaran atau penilaian proses dan hasil belajar, guru sering menggunakan instrumen tertentu, baik tes maupun non-tes (observasi, wawancara, skala sikap, angket dan lain-lain). Instrumen ini mempunyai fungsi dan peran yang sangat penting dalam rangka mengetahui keefektifan proses pembelajaran di sekolah. Pemahaman tentang instrumen ini menjadi penting karena dalam praktik evaluasi dan penilaian, pada umumnya guru selalu

mendasarkan pada proses pengukuran. Dalam pengukuran harus ada alat ukur (instrumen), baik yang berbentuk tes maupun non-tes. Instrumen yang baik adalah instrumen yang memenuhi syarat-syarat atau kaidah tertentu, dapat memberikan data yang akurat sesuai dengan fungsinya, dan hanya mengukur sampel perilaku tertentu (Arifin, 2009).

Untuk mengetahui suatu instrumen atau tes yang digunakan termasuk baik atau kurang baik, maka perlu dilakukan analisis butir soal (*Item analysis*). Dengan analisis soal dapat diperoleh informasi tentang kejelekan sebuah soal dan petunjuk untuk mengadakan perbaikan (Arikunto, 2010). Dalam teori pengukuran, terdapat dua model pengukuran, yaitu *classical test theory* dan *item response theory* (Courville, 2004).

Menurut Hambleton (Gleason, 2008) *classical test theory* telah digunakan selama bertahun-tahun untuk menentukan tingkat kesukaran dan karakteristik lainnya dalam instrumen pengukuran. Namun, ada beberapa kekurangan dalam *classical test theory*. Kekurangan yang paling menonjol dalam teori ini adalah index butir soal seperti tingkat kesukaran dan daya pembeda yang didapatkan dengan menggunakan *classical test theory* bergantung pada kelompok peserta tes dan penilaian kemampuan peserta tes bergantung pada pemilihan butir soal instrumen. Untuk mengatasi permasalahan ini, para ahli psikometri mengembangkan teori pengukuran baru yang disebut *item response theory* (IRT). Model IRT awal digunakan untuk butir soal bentuk betul-salah, model baru telah dikembangkan untuk butir soal bentuk skala sikap, pilihan ganda, dan bentuk lain yang terdapat pada pengembangan instrumen psikologi dan pendidikan.

Hambleton *et al.* (1991) mengemukakan *item response theory* (IRT) adalah teori yang menyatakan bahwa hasil tes dapat diprediksikan atau dijelaskan melalui serangkaian faktor yang disebut dengan sifat atau karakter (*trait*), karakter terpendam (*latent trait*) atau kemampuan (*abilities*) dan hubungan antara jawaban peserta tes dengan kemampuannya dapat dijelaskan dengan grafik karakteristik butir atau *item characteristic curve* (ICC). Semakin tinggi tingkat kemampuan, semakin besar peluang jawaban benar dari suatu butir soal. Sehingga dapat dikatakan bahwa *classical test theory* adalah “*test based*” atau berorientasi pada keseluruhan tes sedangkan IRT merupakan “*item based*” atau lebih berorientasi pada tiap butir soal.

Berdasarkan pemaparan di atas, maka dilakukan penelitian untuk mengembangkan butir soal ulangan kenaikan kelas dengan menerapkan *classical test theory* dan *item response theory* dalam tahap analisis soal. Dengan demikian peneliti akan mengadakan penelitian dengan judul “*Pengembangan dan Analisis Soal Ulangan Kenaikan SMA Kelas X Berdasarkan Classical Test Theory dan Item Response Theory*”.

## **B. Rumusan Masalah**

Berdasarkan latar belakang yang telah dipaparkan, maka rumusan masalah untuk penelitian ini adalah sebagai berikut :

“Apakah bentuk soal ulangan kenaikan kelas SMA Kelas X 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 yang dikembangkan memiliki validitas isi dan validitas empiris (validitas butir soal) yang memenuhi kriteria sebagai butir soal yang baik?
2. Apakah soal ulangan kenaikan kelas yang dikembangkan telah memenuhi kriteria sebagai tes yang baik dilihat dari reliabilitasnya?
3. Apakah soal ulangan kenaikan kelas yang dikembangkan memiliki reliabilitas yang memenuhi kriteria sebagai butir soal yang baik?
4. Apakah soal ulangan kenaikan kelas yang dikembangkan memiliki tingkat kesukaran yang memenuhi kriteria sebagai butir soal yang baik menurut *classical test theory* dan *item response theory*?
5. Apakah soal ulangan kenaikan kelas yang dikembangkan memiliki daya pembeda yang memenuhi kriteria sebagai butir soal yang baik menurut *classical test theory* dan *item response theory*?
6. Apakah soal ulangan kenaikan kelas yang dikembangkan memiliki kualitas pengecoh yang memenuhi kriteria sebagai butir soal yang baik menurut *classical test theory*?
7. Apakah soal ulangan kenaikan kelas yang dikembangkan memiliki faktor tebakan yang memenuhi kriteria sebagai butir soal yang baik menurut *item response theory*?

### C. Batasan Masalah

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

1. Materi yang menjadi bahan soal ulangan kenaikan kelas yang dikembangkan adalah materi yang dipelajari di kelas X SMA pada semester genap antara lain larutan elektrolit dan nonelektrolit, reaksi oksidasi-reduksi, hidrokarbon, minyak bumi dan kegunaan senyawa hidrokarbon dalam kehidupan sehari-hari.
2. Produk evaluasi yang dikembangkan berbentuk tes tertulis tipe pilihan ganda.
3. Pengembangan produk soal dilakukan sampai tahap uji coba terbatas.

### D. Tujuan Penelitian

Penelitian ini bertujuan untuk :

1. Mengembangkan soal ulangan kenaikan kelas yang standar untuk siswa SMA kelas X pada semester genap.
2. Mengetahui apakah soal ulangan kenaikan kelas yang dikembangkan memenuhi kriteria tes yang baik melalui uji reliabilitas, uji validitas, analisis tingkat kesukaran dan daya pembeda berdasarkan *classical test theory* dan *item response theory*.

### E. Manfaat Penelitian

Hasil dari penelitian ini diharapkan dapat memberi manfaat sebagai berikut.

1. Bagi Peserta didik

Dapat benar-benar mengukur kemampuan kognitif peserta didik dan mengasah kemampuan berpikir peserta didik.

## 2. Bagi Guru

Dapat mengetahui tingkat keberhasilan pembelajaran dan menentukan tingkat kemampuan peserta didik berdasarkan instrumen yang telah dikembangkan.

## 3. Bagi Peneliti

Mengetahui kemampuan kognitif peserta didik yang optimal sesuai alat evaluasi yang baik dan menambah wawasan mengenai pengembangan tes berdasarkan *classical test theory* dan *item response theory*.

## **F. Definisi Operasional**

Berikut adalah definisi istilah yang terdapat dalam penelitian ini:

### 1. Pengembangan Tes

Firman (2000) menyatakan bahwa pengembangan tes adalah suatu proses perancangan dan perbaikan alat ukur (tes) agar menjadi suatu alat ukur (tes) yang berkualitas.

### 2. Analisis Kualitas Butir Soal

Analisis soal bertujuan untuk mengadakan identifikasi soal-soal yang baik, kurang baik, dan soal yang jelek (Arikunto, 2010). Menurut Firman (2000) analisis butir soal dilakukan dalam rangka upaya memperbaiki atau meningkatkan kualitas tes yang akan dipakai di masa datang.

### 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-Teori Tes Klasik)

Teori tes klasik menurut O'Connor *et al.* (2002) adalah suatu model pengukuran berdasarkan informasi yang didapatkan pada level skor tes. Menurut Hambleton dan Jones, teori tes klasik adalah teori mengenai skor tes yang mengenalkan tiga konsep yaitu *test score/observed score*, *true score* dan *error score*.

#### 5. *Item Response Theory* (IRT-Teori Respon Butir)

Hambleton *et al.* (1991) mengemukakan bahwa *item response theory* (IRT) adalah teori yang menyatakan bahwa hasil tes dapat diprediksikan atau dijelaskan melalui serangkaian faktor yang disebut dengan sifat atau karakter (*trait*), karakter terpendam (*latent trait*) atau kemampuan (*abilities*) dan hubungan antara jawaban peserta tes dengan kemampuannya dapat dijelaskan dengan grafik karakteristik butir atau *item characteristic curve* (ICC).