

# BAB I

## PENDAHULUAN

### A. Latar Belakang

Ujian Nasional atau yang lebih dikenal dengan UN adalah sistem evaluasi standar pendidikan dasar dan menengah secara nasional dan persamaan mutu tingkat pendidikan antar daerah yang dilakukan oleh Pusat Penilaian Pendidikan Depdiknas di Indonesia. Salah satu tujuan UN adalah untuk melakukan pembinaan dan pemberian bantuan kepada satuan pendidikan dalam upayanya untuk meningkatkan mutu pendidikan. Baswedan (2015) menyatakan bahwa UN harus menjadi kebutuhan siswa dalam mengukur kemampuan hasil belajar yang telah ditempuh. Akan tetapi paradigma hingga saat ini yaitu UN masih diyakini oleh para guru sebagai tujuan akhir dari proses pembelajaran yang ditempuh siswa selama mengikuti proses pembelajaran di sekolah dan masih banyak hal yang harus diperbaiki pada proses perencanaan dan pelaksanaannya. Infografis UN (2015) menyatakan bahwa pada kenyataannya dengan adanya UN siswa mementingkan nilai, kemudian guru dan sekolah fokus pada nilai, bukan kompetensi. Akibatnya hal ini tidak dapat meningkatkan mutu pendidikan pada tingkat sekolah. Salah satu yang dapat dievaluasi dari adanya UN ialah kualitas soal yang digunakan harus dilakukan kajian lebih lanjut oleh para peneliti.

Firman (2003, hlm. 95) menyatakan tes standar untuk instrumen penelitian, seperti UN, kriteria validitas dan reliabilitas mutlak adanya. Selanjutnya menurut Arifin (Nahadi, 2015, hlm. 3) instrumen yang baik adalah instrumen yang dapat memberikan data yang akurat sesuai dengan fungsinya dan hanya mengukur sampel perilaku tertentu. Maka dari itu kualitas soal UN harus memenuhi validitas dan reliabilitas.

Berdasarkan penelitian yang dilakukan oleh Satrisman (2013, hlm. 55). menunjukkan bahwa hanya enam butir soal dengan persentase 15% yang dikategorikan ke dalam soal pada jenjang analisis (C4), selebihnya 22,5% pada jenjang mengingat (C1), memahami (C2), dan 35% pada jenjang

mengaplikasi (C3). Hasil penelitian tersebut membuktikan bahwa soal-soal UN pada jenjang analisis (C4) memiliki persentase yang kecil dibandingkan dengan jenjang mengingat (C1), memahami (C2) dan aplikasi (C3), sehingga soal-soal UN perlu dikaji kembali untuk dapat mengukur kemampuan berpikir tingkat tinggi (*high order thinking skills-HOTS*).

Lailly, dkk. (2015, hlm. 27) dalam penelitiannya menganalisis soal tipe HOTS dalam soal UN kimia SMA. Hasil penelitian menunjukkan bahwa karakteristik soal tipe HOTS yang ada pada soal UN Kimia SMA Rayon B Tahun 2012/2013 yaitu stimulus, sedangkan kemampuan tingkat tinggi belum dapat dihasilkan kesimpulan. Bentuk stimulus dan persentase masing-masing bentuk yang terdapat pada soal UN tersebut yaitu gambar/ grafik/ diagram sebesar 15%, tabel sebesar 15%, simbol/ rumus/ persamaan kimia sebesar 47,5%, contoh sebesar 22,5%, dan penggalan kasus sebesar 32,5%, sehingga keberadaan soal UN mata pelajaran kimia belum mampu mengukur kemampuan berpikir tingkat tinggi.

Baswedan (Republika Online, 2015) berkeinginan untuk mendatangkan tenaga ahli dari luar negeri untuk mendongkrak kemampuan pembuat soal UN. Dengan adanya soal yang bersifat HOTS, peserta UN dituntut untuk mampu berpikir secara logis, kritis, dan analitis, sehingga tidak cukup hanya dengan mengandalkan hafalan saja. Salah satu kemampuan berpikir tingkat tinggi yaitu kemampuan pembuatan keputusan (*decision-making*).

Pembuatan keputusan merupakan suatu bentuk penyelesaian akhir dari masalah, karena setelah mengambil keputusan maka muncul suatu jawaban dari permasalahan yang sedang dihadapi. Sebelum dilakukan pengambilan keputusan ada beberapa hal yang harus dilakukan sebelumnya seperti menganalisis masalah atau menentukan alternatif-alternatif pilihan. Hal ini sejalan dengan Santrock (Liliasari, 2013, hlm. 125) mengambil keputusan adalah sebuah pemikiran dimana individu mengevaluasi berbagai pilihan dan memutuskan pilihan dari sekian banyak pilihan. Maka dari uraian tersebut dapat disimpulkan bahwa pengambilan keputusan merupakan proses berpikir untuk mengidentifikasi dan memutuskan pilihan dari berbagai pilihan yang ada. Hal ini dapat menuntut siswa untuk mengambil keputusan yang terbaik

dan paling efektif untuk mengatasi permasalahan yang ada. Salah satu kompetensi yang ditekankan dalam kurikulum 2013 adalah kemampuan pembuatan keputusan.

Kemampuan pembuatan keputusan merupakan salah satu aspek Standar Kompetensi Lulusan (SKL) pada kurikulum 2013 yang harus dicapai oleh siswa. Hal ini sesuai dengan Peraturan Menteri Pendidikan Nasional (2013) yang menyatakan bahwa Standar Kompetensi Lulusan SMA/MA/SMALB/Paket C dan SMK/MAK yaitu KI-3: memahami, menerapkan, dan menganalisis pengetahuan faktual, konseptual, prosedural, dan metakognitif berdasarkan rasa ingin tahunya tentang ilmu pengetahuan, teknologi, seni, budaya, dan humaniora dengan wawasan kemanusiaan, kebangsaan, kenegaraan, dan peradaban terkait penyebab fenomena dan kejadian, serta menerapkan pengetahuan prosedural pada bidang kajian yang spesifik sesuai dengan bakat dan minatnya untuk memecahkan masalah dan membuat keputusan.

Kemampuan pembuatan keputusan sangat diperlukan dalam kehidupan sehari-hari, sehingga pada KI-2 juga ditekankan urgensi pembuatan keputusan yaitu menunjukkan perilaku responsif dan proaktif serta bijaksana sebagai wujud kemampuan memecahkan masalah dan membuat keputusan, sehingga pembuatan keputusan merupakan salah satu kemampuan berpikir tingkat tinggi yang harus dimiliki oleh siswa. Pembuatan keputusan yang tepat dapat dilakukan apabila siswa mampu memahami konsep suatu materi secara utuh. Siswa dikatakan dapat memahami konsep kimia secara menyeluruh dan mendalam jika siswa mampu menghubungkan ketiga level representasi kimia (Sunyono, 2013).

Representasi kimia dibagi ke dalam tiga tingkatan. Representasi level makroskopis didapatkan melalui observasi dengan menggunakan panca indera, pengetahuan simbolik menunjukkan simbol dan persamaan reaksi dan pengetahuan submikroskopis meliputi penggambaran atom, molekul, dan ion menggunakan gambar atau model molekul (Chandrasegaran, dkk., 2007, hlm. 294). Dengan demikian, seorang siswa dapat dikatakan telah memahami suatu konsep kimia secara utuh ketika ia mampu mentransformasikan pengetahuan

level makroskopisnya ke level submikroskopis ataupun simbolis. Namun, fakta di lapangan sebagian besar hanya mencakup makroskopis dan simbolik saja (Sunyono, 2013, hlm. 66). Hal ini membuat siswa menjadi kurang terlatih untuk menghubungkan pengetahuan makroskopik dan simbolik dengan pengetahuan submikroskopik. Salah satu materi kimia yang memiliki konsep yang kompleks yaitu larutan penyangga.

Materi larutan penyangga mempelajari tentang pengertian larutan penyangga, pH larutan penyangga, dan fungsi larutan penyangga dalam makhluk hidup dan kehidupan sehari-hari. Materi pokok larutan penyangga merupakan konsep yang tidak cukup dihafal saja namun terdapat konsep-konsep yang perlu diobservasi. Konsep tersebut dijelaskan dalam berbagai macam bentuk representasi yang dapat memvisualisasikan materi-materi tersebut sehingga diharapkan siswa dapat mengamati gejala-gejala yang terjadi, dan menganalisa serta menarik kesimpulan sehingga akan diperoleh konsep-konsep yang bersifat bukan hanya hafalan saja. Kimia di Indonesia umumnya masih menggunakan pendekatan tradisional, yaitu siswa dituntut lebih banyak mempelajari konsep-konsep dan prinsip-prinsip sains secara verbalistik (Liliasari dalam Sunyono, 2012, hlm. 1)

Soal-soal Ujian Nasional selama ini bersifat (konvensional) artinya ujian dilakukan dengan menggunakan kertas dan pensil dengan istilah sekarang *Paper Based Test* (PBT). PBT yang dilakukan saat ini banyak masalah/kendala seperti: 1). Bentuk soal yang digunakan sulit untuk dibuat bervariasi; 2). Tampilan soal terbatas hanya dua dimensi; 3). Kerahasiaan tes tidak dijamin karena dapat saja dibaca oleh orang tidak berwenang; 4). Waktu penyelenggaraan yang lebih lama; 5). Membutuhkan ruang untuk penyimpanan data perangkat tes; 6). Kecurangan selama pelaksanaan ujian; 7). Perlu langkah *scanning* LJK dan *scoring* sehingga pengolahan hasil memerlukan waktu yang relatif lama; 8). Membutuhkan biaya banyak. (Santoso, 2010; Puspendik, 2015). Oleh karena itu ujian dengan PBT kurang efektif dan efisien.

Salah satu alternatif yang dapat dilakukan untuk mengatasi permasalahan tersebut adalah dengan memanfaatkan teknologi informasi dan komunikasi (TIK) yang saat ini sedang berkembang dalam penilaian pendidikan. Bentuk pemanfaatannya adalah dengan menerapkan bentuk ujian berbasis komputer. Pemanfaatan komputer sebagai alat evaluasi sangat memungkinkan terlaksananya proses evaluasi yang efektif dan efisien. Selain itu pemanfaatan komputer sebagai media dalam pelaksanaan ujian merupakan upaya untuk membiasakan siswa berinteraksi dengan teknologi.

Penyelenggaraan Ujian Nasional Berbasis Komputer (UNBK) bukan hanya sebagai proses digitalisasi UN akan tetap harus meningkatkan kualitas soal yang akan diujikan. Penyelenggaraan UNBK pertama kali dilaksanakan pada tahun 2014 di Sekolah Indonesia Luar Negeri (SILN) pada jenjang SMP di Singapura dan Malaysia. Pada tahun 2015 dilaksanakan rintisan UNBK dengan mengikutsertakan sebanyak 555 sekolah yang terdiri dari 42 SMP/MTs, 135 SMA/MA, dan 378 SMK di 29 Provinsi dan Luar Negeri (Kemdikbud, 2015). Tahun 2019-2020 UN direncanakan secara luas dilaksanakan berbasis komputer (infografis UN, 2015). Kasim (Sindonews, 2015) mengatakan anggaran UN akan dihemat sebesar 50% dari anggaran UN tahun ini Rp. 580 miliar. Penghematan terjadi karena tidak ada pencetakan naskah soal dan lembar jawaban, serta tidak ada pengawasan distribusi soal dan lembar jawaban.

Menurut Lowry (2005, hlm. 198) Pemanfaatan *Computer Aided Assessment* (CAA) dapat menimbulkan dampak positif pada pengalaman belajar siswa. Data menunjukkan bahwa penggunaan CAA sebagai asesmen mandiri berdampak positif pada siswa. Bagi guru, CAA sangat berperan dalam menghemat waktu. Sementara bagi siswa, sistem CAA ini dapat memberikan beberapa keuntungan. (1) memberikan umpan balik pada siswa, (2) membimbing usaha-usaha siswa, (3) mendiagnosa permasalahan dalam pembelajaran, dan (4) memberikan pengalaman pada siswa dalam kegiatan asesmen mandiri. Hal ini juga didukung oleh hasil penelitian Bull (1999, hlm. 125) mengenai dampak penggunaan CAA di perguruan tinggi, yang berpotensi menghasilkan berbagai format dan jenis asesmen, menghasilkan

umpan balik yang cepat sehingga dapat meningkatkan kapabilitas pengajar dalam menunjukkan kemajuan serta kekurangan siswa dalam pembelajaran.

Menurut Bodmann and Robinson (2004), *Computer-Based Tests* (CBT) menawarkan beberapa keuntungan dibandingkan dengan PBT. Penggunaan teknologi dalam asesmen dapat mengakomodasi proses penilaian pengetahuan yang bersifat kompleks yang tidak dapat diakomodasi dengan PBT. Hasil penelitian Karadeniz (2009) menyatakan bahwa siswa yang diberikan asesmen dalam bentuk *web based* dan *mobile based assessment* menunjukkan respon yang positif karena kedua bentuk asesmen tersebut memberikan kemudahan dalam penggunaannya, komprehensif serta umpan balik yang cepat jika dibandingkan dengan PBT.

Penerapan CBT di Universitas Luton membantu pihak Universitas untuk menjaga kualitas pembelajaran disaat jumlah mahasiswa meningkat namun sumber daya menurun. Selain itu, penggunaan CBT juga meningkatkan respon siswa, mengurangi suasana tegang dan menciptakan suasana tenang sehingga dapat meningkatkan konsentrasi peserta ujian (Zakrzewski dan Bull, 1998). Secara umum, penerapan sistem CBT dapat menghemat waktu, mengurangi beban pengajar, hasil penilaian juga lebih cepat dan dapat mengurangi kecurangan yang mungkin dilakukan siswa saat ujian (Bodmann dan Robinson 2004; Jamil, dkk., 2012, hlm. 274; Mason, dkk., 2001, hlm. 29-30; Terzis dan Economides, 2011, hlm. 1032; Thelwall, 2000, hal. 40). Jenis asesmen ini juga mendapatkan respon yang baik dari para siswa (Jawaid, dkk., 2014, hlm. 690; Terzis dan Economides, 2011, hlm. 1032).

Jenis asesmen yang berbasis komputer memungkinkan soal-soal ujian disajikan bukan hanya teks saja tetapi juga dapat menggabungkan grafis dan multimedia (Thelwall, 2000; Jawaid, dkk., 2014, hlm. 688). Hal ini menjadikan jenis asesmen ini cocok digunakan dalam penilaian pemahaman siswa dalam pembelajaran kimia dengan pendekatan representasional. Penilaian berbasis komputer memungkinkan pengukuran pemahaman pada ketiga level representasi karena dapat memuat video dan animasi pada soal yang diujikan (Bowen, 1998). Hal ini telah dimanfaatkan oleh Scherer, dkk. (2014) untuk mengembangkan asesmen berbasis komputer yang mengukur

empat tahapan *problem solving* siswa dalam menyelesaikan soal-soal kimia yang kompleks. Asesmen tersebut memenuhi kriteria sebagai asesmen yang valid dan dapat digunakan dalam penilaian kemampuan *problem solving* siswa pada kelas sains.

Melihat fenomena yang telah diuraikan, penilaian hanya mengukur kemampuan berpikir tingkat rendah (*low order thinking skills*-LOTS) dan media yang digunakan berupa PBT sehingga *multiple representasi* kimia tidak dapat diperlihatkan secara utuh. Permasalahan seperti pernyataan pokok uji yang mengandung konsep-konsep yang abstrak dan sulit dideskripsikan dengan kata-kata menyebabkan siswa akan mengalami kesulitan memahami pernyataan pokok uji. Penggunaan *virtual test* dapat membuat jenis pertanyaan lebih interaktif. Dengan menggunakan *virtual test* memungkinkan untuk pembuatan soal menggunakan gambar, grafik, animasi, dan video dalam pembuatan soal sehingga dapat memperjelas maksud dari pernyataan pokok uji. Oleh karena itu, perlu dikembangkan *virtual test* berbasis *multiple representasi* yang dapat mengukur kemampuan pembuatan keputusan yang berjudul “*Pengembangan dan Validasi Virtual Test Berbasis Multiple Representasi untuk Mengukur Kemampuan Pembuatan Keputusan (Decision-Making) pada Materi Larutan Penyangga*”.

## **B. Identifikasi Masalah Penelitian**

Berdasarkan uraian yang telah dikemukakan pada latar belakang, bahwa memiliki komitmen untuk memperbaiki kualitas soal UN, yang biasanya mengukur kemampuan berpikir tingkat rendah (*low order thinking skills*-LOTS) menjadi lebih menekankan pada kemampuan berpikir tingkat tinggi rendah (*high order thinking skills*-HOTS) seperti kemampuan pembuatan keputusan. Akan tetapi, kesulitan yang dialami siswa yaitu dalam memahami pernyataan pokok uji (*stem*) yang mengukur kemampuan berpikir tingkat tinggi karena pada umumnya pernyataan pokok ujinya lebih panjang dibandingkan pernyataan pokok uji yang mengukur kemampuan berpikir tingkat rendah. Selain itu konsep-konsep abstrak yang membutuhkan pemahaman yang komprehensif yaitu dengan mentransformasikan

pengetahuan level makroskopisnya ke level submikroskopis ataupun simbolis dan itu kurang ditonjolkan pada pernyataan pokok uji. Oleh karena itu, perlu dikembangkan *virtual test* berbasis *multiple representasi* yang dapat mengukur kemampuan pembuatan keputusan siswa, dimana hal ini juga sejalan dengan adanya komitmen pemerintah untuk melaksanakan UN dengan CBT baik *online* maupun *offline* (Kasim, 2015).

### C. Rumusan Masalah

Berdasarkan latar belakang masalah yang diuraikan di atas, maka secara umum rumusan masalah dari penelitian ini adalah “*Apakah virtual test berbasis multiple representasi yang dikembangkan untuk mengukur kemampuan pembuatan keputusan siswa (decision-making) pada materi larutan penyangga telah memenuhi validitas dan reliabilitas?*”

Untuk memperjelas rumusan masalah, maka difokuskan pada beberapa pertanyaan penelitian sebagai berikut:

1. Bagaimanakah validitas konten *virtual test* yang dikembangkan untuk mengukur kemampuan pembuatan keputusan?
2. Bagaimanakah reliabilitas *virtual test* yang dikembangkan untuk mengukur kemampuan pembuatan keputusan?
3. Bagaimanakah tingkat kesukaran soal *virtual test* untuk mengukur kemampuan pembuatan keputusan?
4. Bagaimanakah daya pembeda soal *virtual test* untuk mengukur kemampuan pembuatan keputusan?
5. Bagaimanakah tingkat keterbacaan soal *virtual test* dibandingkan dengan soal *paper and pencil test* untuk mengukur kemampuan pembuatan keputusan?
6. Bagaimanakah tanggapan guru dan siswa pada *virtual test* yang digunakan untuk mengukur kemampuan pembuatan keputusan?

#### D. Tujuan Penelitian

Berdasarkan latar belakang masalah dan rumusan masalah yang diuraikan di atas maka tujuan penelitian ini adalah untuk mendapatkan *virtual test* berbasis *multiple representasi* yang sudah valid dan reliabel untuk mengukur kemampuan pembuatan keputusan siswa (*decision-making*) pada materi larutan penyangga. Untuk memperjelas tujuan penelitian maka diuraikan menjadi:

1. Menganalisis validitas konten *virtual test* yang dikembangkan untuk mengukur kemampuan pembuatan keputusan.
2. Menganalisis reliabilitas *virtual test* yang dikembangkan untuk mengukur kemampuan pembuatan keputusan.
3. Menganalisis tingkat kesukaran soal *virtual test* untuk mengukur kemampuan pembuatan keputusan.
4. Menganalisis daya pembeda soal *virtual test* untuk mengukur kemampuan pembuatan keputusan.
5. Menganalisis tingkat keterbacaan soal *virtual test* dibandingkan dengan soal *paper and pencil test* untuk mengukur kemampuan pembuatan keputusan.
6. Menganalisis tanggapan guru dan siswa pada *virtual test* yang digunakan untuk mengukur kemampuan pembuatan keputusan.

#### E. Pembatasan Masalah

Agar penelitian lebih terarah dan memberi gambaran yang jelas, maka batasan masalah dalam penelitian ini sebagai berikut:

1. Pengambilan keputusan yang digunakan dalam penelitian ini diadopsi dari Bavolar (2013, hlm. 117). Mencakup empat kompetensi yang kemudian di turunkan menjadi enam indikator. Enam indikator tersebut diantaranya menganalisis adanya beberapa alternatif jawaban yang mungkin beserta resiko yang mungkin muncul (*concistency in risk perception*), menganalisis keterkaitan permasalahan dengan aturan-aturan atau konsep-konsep yang ada (*recognizing norm*), mendeteksi kesalahan penyusunan jawaban (*sesistance to framing*), memahami dasar

pembuatan keputusan yang tidak relevan (*resistance to sunk cost*), mengintegrasikan keyakinan dan nilai yang berkaitan dan mengevaluasi diri sendiri.

2. Pengembangan *virtual test* yang dikembangkan pada penelitian ini adalah soal pilihan ganda berdasarkan enam indikator pengambilan keputusan yang dikembangkan oleh Bavolar dan *virtual test* yang dikembangkan masih berupa *software (offline)*.
3. Instrumen penilaian yang dikembangkan pada materi larutan penyangga kelas XI yang mengacu pada kurikulum 2013, kompetensi dasar 3.13. yaitu menganalisis peran larutan penyangga dalam tubuh makhluk hidup (Permendikbud, 2013, hlm.169).

#### **F. Manfaat Penelitian**

Hasil dari penelitian pengembangan instrumen ini diharapkan dapat memberikan manfaat untuk peningkatan kualitas pendidikan. Manfaat dari segi praktik yang diharapkan dari penelitian ini adalah:

1. Bagi siswa, terlatihnya pengetahuan dan kemampuan pengambilan keputusan siswa SMA khususnya pada materi larutan penyangga.
2. Bagi guru, hasil dari penelitian ini mampu memberikan informasi terkait pengembangan soal pilihan ganda *virtual test* berbasis *multiple representasi* yang dapat mengukur pengambilan keputusan siswa SMA.
3. Bagi peneliti, memperoleh pengalaman langsung dalam mengembangkan instrumen penilaian dalam bentuk pilihan ganda *virtual test* berbasis *multiple representasi* yang dapat mengukur pengambilan keputusan siswa SMA.
4. Bagi peneliti lain, hasil penelitian ini dapat dijadikan sebagai referensi untuk melakukan penelitian lebih jauh terkait penilaian pilihan ganda *virtual test* berbasis *multiple representasi*.

## G. Definisi Operasional

Untuk menghindari kesalahan penafsiran terhadap istilah-istilah yang digunakan dalam penelitian ini, maka diberikan definisi operasional mengenai istilah-istilah sebagai berikut:

1. *Virtual Test* yaitu tes menggunakan *software* (perangkat lunak) yang dapat dilaksanakan baik secara *online* atau *offline* (Firman, 2014, hlm. 12).
2. Pilihan ganda adalah soal yang jawabannya harus dipilih dari beberapa kemungkinan jawaban yang telah disediakan. Pokok pilihan berganda adalah pokok uji yang terdiri dari suatu pernyataan yang belum lengkap, dan untuk melengkapi pernyataan itu disediakan beberapa pernyataan sambungan, salah satu diantaranya merupakan jawaban yang benar (Firman, 2013, hlm. 27).
3. Validitas adalah suatu alat ukur yang menunjukkan sejauh mana ketetapan dan kecermatan suatu alat ukur dalam melakukan fungsi ukurnya.
4. Reliabilitas adalah ukuran sejauh mana suatu alat ukur memberikan gambaran yang benar-benar dapat dipercaya tentang kemampuan seseorang.
5. Pengambilan keputusan adalah sebuah pemikiran dimana individu mengevaluasi berbagai pilihan dan memutuskan pilihan dari sekian banyak pilihan (Santrock dalam Liliyasi, 2013, hlm. 125). Pengambilan keputusan yang baik mengidentifikasi keuntungan dan resiko dari setiap pilihan yang ada, menggunakan setiap bukti yang tersedia untuk menentukan bobot tiap pilihan secara logis dan kemudian memutuskannya.
6. *Multiple representasi* adalah praktik merepresentasikan kembali (*re-presenting*) konsep yang sama melalui berbagai bentuk, yang mencakup mode-mode representasi deskriptif (verbal, grafik, tabel) (Waldrup dalam Sjaeful, 2014, hlm. 41).