

BAB V

SIMPULAN DAN SARAN

A. Simpulan

Penelitian ini telah menghasilkan instrumen tes diagnostik yang dikembangkan dengan menggunakan model DINA. Selain itu, dikembangkan pula *software* komputer dalam bentuk *Computer Based Test* (CBT), untuk memberikan umpan balik (*feed back*) kepada siswa. *Software* CBT tersebut berbasis *website* sehingga dapat diakses melalui jaringan internet maupun intranet. *Software* CBT juga dilengkapi dengan pedoman penggunaan (*user manual*), sehingga mempermudah pengguna dalam memanfaatkan berbagai fasilitas yang ada dalam *software*.

Selanjutnya, berdasarkan hasil penelitian dan pembahasan yang dikemukakan pada bab sebelumnya, maka dapat dikemukakan temuan penelitian sebagai berikut:

1. Tahapan pengembangan tes dengan menggunakan model DINA, yang digunakan dalam penelitian ini meliputi: identifikasi kompetensi dasar (KD) dan merumuskan indikator, menyusun *learning continuum*, menyusun hierarki materi, merumuskan atribut, mengonstruksi soal, validasi ahli dan uji empirik. Setelah melalui tujuh tahapan tersebut, berhasil dikembangkan 37 item tes diagnostik.
2. Dari 37 item tes diagnostik yang dikembangkan, 15 item di antaranya harus dihilangkan/dihapus dari paket tes. Item-item yang dihilangkan/dihapus

disebabkan tidak memenuhi uji fit model dan kualitas item tersebut jelek karena memiliki indeks daya beda kurang dari 0,2.

3. Pemanfaatan informasi hasil tes diagnostik yang dikembangkan dilakukan melalui umpan balik yang dihasilkan *software* CBT. Umpan balik tersebut berbentuk laporan skor per item dan laporan diagnostik masing-masing anak/peserta tes.
4. Hasil transaksi secara individual masing-masing anak berupa pola respon, akan terekam pada *software* CBT. Akumulasi pola respon anak dalam satu kelas, dapat dimanfaatkan untuk analisis salah konsepsi secara klasikal dengan program Mplus, BIGSTEP, ITEMAN atau *software* analisis lainnya.
5. Hasil studi simulasi menunjukkan: (i) matriks Q berukuran 4×3 ($Q_{4 \times 3}$) lebih ideal dibandingkan dengan matriks Q berukuran 6×3 ($Q_{6 \times 3}$) dan 8×3 ($Q_{8 \times 3}$); (ii) matriks Q berukuran 6×4 ($Q_{6 \times 4}$) lebih ideal dibandingkan dengan matriks Q berukuran 4×4 ($Q_{4 \times 4}$) dan 8×4 ($Q_{8 \times 4}$); serta (iii) matriks Q berukuran 6×5 ($Q_{6 \times 5}$) lebih ideal dibandingkan dengan matriks Q berukuran 4×5 ($Q_{4 \times 5}$) dan 8×5 ($Q_{8 \times 5}$). Suatu matriks dikatakan ideal bila matriks tersebut mampu menghasilkan struktur *latent class* yang maksimum.
6. Hasil simulasi juga menunjukkan: ukuran sampel minimum ($N = 250$) mampu mengestimasi parameter-parameter yang ada pada model DINA dengan baik dibandingkan sampel dengan ukuran yang lebih besar ($N = 500$ dan $N = 1000$). Artinya, parameter *slip* dan *guessing* pada sampel minimum cenderung lebih rendah dibandingkan dua ukuran sampel lainnya.

B. Keterbatasan Penelitian

Penelitian ini telah dirancang dan dilakukan secara cermat, hati-hati dan sungguh-sungguh dengan melibatkan banyak pihak untuk menelaah hasil pengembangan perangkat instrumen. Tahapan juga dilaksanakan secara ketat, sesuai saran dari promotor/kopromotor, reviewer dan berbagai pihak agar dihasilkan perangkat tes diagnostik yang berkualitas. Namun, penelitian ini tidak terlepas dari beberapa keterbatasan, karena pengembangan tes diagnostik berbasis pada *latent class* (model DINA) merupakan hal baru dalam dunia pengujian pendidikan di Indonesia. Beberapa kelemahan dalam penelitian ini di antaranya:

1. Pelaksanaan uji coba penelitian ini baru dikenakan pada sejumlah siswa kelas VIII SMP negeri, dan dilakukan di akhir semester gasal. Kondisi ini tentunya akan berbeda bila uji coba juga melibatkan siswa SMP swasta, dan dikenakan pada siswa kelas VIII yang sedang berada di awal semester gasal atau siswa kelas VII di akhir semester genap. Dengan demikian, dari aspek validitas internal masih mengandung beberapa keterbatasan, baik menyangkut karakteristik subjek penelitian, waktu dan tempat penelitiannya.
2. Cakupan materi pada tes diagnostik ini terlalu luas, yakni semua materi aljabar yang diajarkan di kelas VII SMP (ada 5 materi). Luasnya cakupan materi tes berdampak pada kompleksnya atribut yang mendasari setiap item. Akibatnya, matriks Q yang terbentuk memiliki atribut lebih banyak, dibandingkan jumlah item.

3. Penyusunan item tes dilakukan dengan mengacu pada kisi-kisi tes seperti lazimnya pada proses pengembangan tes prestasi. Hal itu berbeda dengan langkah yang seharusnya digunakan pada model DINA, yaitu item tes dibangun dan diturunkan berdasarkan sejumlah atribut-atribut.
4. Tes diagnostik dan program CBT yang dikembangkan dalam penelitian masih terbatas pada tes bentuk pilihan ganda, belum menjangkau pada tes uraian.
5. *Software* CBT yang dikembangkan hanya dirancang untuk menyajikan umpan balik berdasarkan jawaban peserta tes secara individual. *Software* CBT belum terintegrasi dengan satu rangkaian *out put* Mplus dan program R. Dengan demikian, informasi salah konsepsi secara klasikal belum mampu dihasilkan oleh *software* CBT. Informasi salah konsepsi secara klasikal masih harus diperoleh dengan bantuan program Mplus, dan dilakukan secara terpisah.

C. Implikasi Hasil penelitian

Hasil-hasil penelitian di atas mengandung beberapa implikasi bagi penerapan tes diagnostik yang dikembangkan penelitian ini, antara lain:

1. Dengan dihasilkannya produk penelitian ini, kepada pihak sekolah perlu mempertimbangkan untuk menggunakannya. Penggunaan *software* CBT akan memberikan umpan balik kepada setiap siswa secara cepat dan tentunya lebih objektif.

2. Informasi salah konsepsi yang dihasilkan *software* CBT, hanya terbatas pada salah konsepsi yang dilakukan masing-masing anak. Agar diperoleh informasi salah konsepsi yang lebih komprehensif dan secara klasikal, sangat tepat dilakukan analisis dengan menggunakan bantuan *software* analisis seperti Mplus. Perpaduan dua informasi tersebut, akan mempermudah guru mendeteksi salah konsepsi yang dialami oleh anak didiknya.
3. Pengembangan tes diagnostik dengan model DINA juga lebih mudah dibandingkan dengan pengembangan tes diagnostik yang berbasis pada teori tes klasik maupun teori respon butir. Pada kedua teori tersebut, proses pengembangan tes harus melalui proses membuat analisis pengecoh, menetapkan skor peserta tes, dan menentukan skor batas. Namun, pada model DINA ketiga langkah itu tidak perlu dilakukan. Oleh karena itu, penggunaan model DINA untuk pengembangan tes diagnostik praktis diterapkan di lapangan.
4. Selama proses membangun item tes pada model DINA, cacah item harus disesuaikan dengan banyaknya atribut. Bila hal itu diabaikan, maka akan mempengaruhi fit tidaknya dengan model sehingga dapat mempengaruhi kesimpulan diagnosis.

D. Saran-saran

Berdasarkan simpulan penelitian maka dapat diajukan beberapa saran sebagai berikut:

1. Hasil penelitian ini semakin memperkuat teori-teori pada tes diagnostik, yaitu tes diagnostik berfungsi untuk mengukur kemampuan siswa yang

spesifik. Oleh karena itu, kepada guru yang ingin mendesain tes diagnostik disarankan agar fokus pada suatu topik tertentu. Untuk memastikan bahwa kompetensi tertentu gagal dikuasai oleh anak, maka kompetensi itu harus diuji melalui beberapa item tes diagnostik. Artinya, kurang tepat apabila seorang anak diasumsikan tidak kompeten pada suatu materi tertentu hanya disebabkan karena mereka gagal menjawab sebuah item tes.

2. Penelitian ini juga menghasilkan *software* CBT yang dapat digunakan dengan mudah mendeteksi penguasaan anak dalam aljabar. Dengan demikian, kepada pihak sekolah direkomendasikan untuk menggunakan *software* CBT tersebut. Namun, disarankan agar panduan penggunaan *software* CBT dicermati secara baik agar informasi diagnostik yang dihasilkan dapat ditafsirkan secara benar.
3. Kepada para peneliti yang berminat melakukan penelitian di bidang diagnostik *latent class*, disarankan untuk melakukan penelitian lanjutan secara lebih mendalam seperti: (i) membandingkan dengan model *latent class* lain seperti dengan model NIDA atau model NC-RUM (*Noncompensatory Reparametrized Unified Model*). Hal ini sangat penting untuk menguji apakah faktor yang terjadi pada proses pengembangan tes dengan model DINA juga memiliki dampak sama terhadap hasil estimasi parameter; (ii) mensimulasikan berbagai ukuran matriks juga penting dilakukan. Misalnya panjang item dibuat lebih variatif yakni dengan panjang 3, 5, 7, 9, 11 atau yang lain dan banyaknya atribut 2, 6, 7, 8. Dengan simulasi yang lebih variatif, informasi yang didapatkan juga semakin lengkap,

sehingga dapat melengkapi hasil penelitian ini; (iii) mengintegrasikan hasil analisis model DINA dengan analisis secara klasik dalam bentuk *software* komputer untuk menyediakan informasi diagnostik juga merupakan bidang penelitian berikutnya yang menarik dilakukan. Dengan dilakukannya penelitian seperti ini, informasi yang didapatkan menjadi lebih kaya sehingga dapat menjawab dua pertanyaan sekaligus: (a) mengapa peserta tes tidak menguasai suatu materi? dan (b) siapa peserta tes yang tidak menguasai suatu materi?; dan (iv) menganalisis interaksi antar *latent class* pada model DINA merupakan bidang garapan penelitian berikutnya yang dapat dilakukan. Dengan melakukan langkah ini, informasi diagnostik yang diperoleh dari *latent class* lebih kaya dibandingkan dengan hasil diagnosis dengan model CTT maupun IRT. Membandingkan interaksi antar *latent class* antara model DINA, NIDA atau NC-RUM juga lahan penelitian berikutnya yang layak untuk dilakukan.