

BAB V

KESIMPULAN, IMPLIKASI DAN SARAN

5.1. Kesimpulan

Penelitian mengenai pengembangan instrumen tes tertulis untuk mengukur keterampilan *STEM* pada konsep IPA telah mengikuti tahapan dalam metode penelitian pengembangan. Berdasarkan penjelasan analisis pada hasil dan pembahasan penelitian, maka dapat ditarik kesimpulan diantaranya adalah:

1. *Framework* asesmen *STEM* yang dapat digunakan untuk mengukur keterampilan *STEM* siswa SMP pada Konsep IPA adalah; a) Keterampilan menganalisis fenomena saintifik (keterampilan meneliti, mengembangkan rencana dan timeline, menganalisis data dan menyusun kesimpulan dari hasil analisis), b) Keterampilan memecahkan masalah (keterampilan untuk memahami sistem kompleks dan diagram ke dalam bagian kecil, memproses dan mengidentifikasi hubungan sebab akibat, dan mengajukan solusi untuk kemajuan proyek), c) Keterampilan matematis (keterampilan untuk mengkalkulasi, menghitung dan menganalisis data matematis), d) Keterampilan engineering design dan menggunakan teknologi (keterampilan dalam menyeleksi software dan alat teknis dalam menyelesaikan masalah/proyek). e) Keterampilan berpikir kreatif (keterampilan untuk menyelesaikan masalah, mendesain produk dan gagasan baru dan mengkomunikasikan penemuannya dengan efektif), f) Keterampilan mengorganisasi data dan bukti saintifik (keterampilan untuk mengumpulkan dan membedakan validitas informasi, merekam data akurat dan menyelesaikan tugas sesuai waktu yang ditentukan).
2. Penyusunan dan pengembangan desain instrumen tes tertulis untuk mengukur keterampilan *STEM* telah memenuhi syarat sebagai desain yang baik yang memuat keterampilan dasar *STEM* pada butir soal. Desain instrumen ini dikembangkan setelah melalui tahap kajian pustaka, tahap kajian analisis kurikulum 2013 untuk menentukan konsep IPA yang digunakan, tahap penyusunan dan pengembangan *framework* keterampilan *STEM*, tahap penyusunan dan pengembangan *test blueprint* dan tahap penyusunan dan

pengembangan instrumen tes tertulis pada tahap uji terbatas dan uji luas, kemudian analisis pemodelan *Rasch*.

3. Kualitas *test blueprint* (kisi – kisi) yang baik akan menghasilkan butir soal dengan karakteristik yang baik dan dapat mengukur keterampilan *STEM* secara akurat. *Test blueprint* ini telah melalui validasi dan telaah keterbacaan oleh dosen dan guru ahli untuk menilai dari aspek substansi, aspek konstruksi, dan aspek bahasa. Berdasarkan tahap validasi dan telaah keterbacaan ini, *test blueprint* disimpulkan layak untuk diimplementasikan pada tahap uji terbatas dan tahap uji luas. Angket guru juga digunakan sebagai *preliminary study* (studi tahap awal) untuk memastikan bahwa sekolah yang dijadikan responden telah menerapkan pembelajaran berbasis *STEM* agar instrumen tes yang diteliti tepat sasaran. Adapun karakteristik butir soal berbasis *STEM* yang membedakan dari tes lainnya adalah: Instrumen tes berbasis *STEM* mencakup konsep IPA yang konkret dan dekat dengan keseharian siswa, Instrumen mencakup proses pikir, desain, buat dan uji untuk mengukur keterampilan dasar *engineering design* dan menggunakan teknologi.
4. Secara keseluruhan, ketiga paket soal yang telah disusun dan dianalisis berdasarkan Pemodelan *Rasch* merupakan instrumen tes tertulis yang valid, reliabel dan memiliki nilai diskriminasi yang sangat baik. Paket soal Energi dalam Sistem Kehidupan yang berjumlah 31 butir soal disimpulkan diterima dengan kriteria nilai yang sangat baik. Reliabilitas butir soal dari paket soal ini adalah 0.94, sedangkan reliabilitas individunya adalah 0.81 dengan beberapa soal yang terdeteksi bias namun tetap dapat digunakan. Paket soal Tekanan Zat Cair, Padat dan Gas dan Tekanan pada Tumbuhan yang berjumlah 31 butir soal mempunyai validitas yang baik dengan nilai reliabilitas butir soal sebesar 0.92 dan reliabilitas individu senilai 0.87. Paket soal ini disimpulkan untuk menolak 2 butir soal yaitu S1 dan S2 karena ketiga kriteria outfit MNSQ, outfit ZSTD dan poin biserial yang sangat tinggi di luar rentang batas yang seharusnya. Lalu pada paket soal yang ketiga yaitu Gunung Api dan Gempa Bumi, validitas dan reliabilitasnya relatif sangat baik dengan reliabilitas butir soal senilai 0.95, dan

reliabilitas individu senilai 0.81. Dari sebanyak 32 soal pada paket soal ini, sebanyak 3 butir soal tidak valid sehingga harus ditolak, 4 butir soal dinyatakan valid dengan revisi, dan sebanyak 25 butir soal dinyatakan valid dan memadai serta dapat digunakan.

5.2. Implikasi

Pengembangan instrumen tes tertulis untuk mengukur keterampilan *STEM* merupakan instrumen tes yang valid, reliabel dan akurat untuk digunakan dalam evaluasi pembelajaran berbasis *STEM* di Sekolah Menengah Pertama. Instrumen tes tertulis berbasis *STEM* ini telah dibuktikan dapat mengukur keterampilan *STEM* secara akurat karena setiap deskriptor yang mewakili 6 keterampilan *STEM* dituangkan ke dalam setiap butir soal dengan sangat baik. Pengembangan instrumen tes tertulis ini telah menjalani proses validasi dan telaah keterbacaan dari dosen ahli dan guru SMP dan SMA, kemudian menjalani revisi berkali – kali sampai kemudian layak untuk digunakan dalam penelitian pada tahap uji terbatas, revisi awal, lalu tahap uji luas, analisis pemodelan *Rasch* untuk kemudian kembali menjalani revisi tahap akhir. Hasil revisi tahap akhir berdasarkan analisis pemodelan *Rasch* ini dapat disebut bank soal yang telah teruji baik validitas dan reliabilitas dalam mengukur keterampilan *STEM* yang dimiliki oleh siswa. Maka dari itu, instrumen tes tertulis untuk mengukur keterampilan *STEM* ini dinilai layak dan disarankan untuk digunakan oleh guru atau praktisi pendidikan dalam proses asesmen di sekolah – sekolah agar keterampilan *STEM* siswa dapat terukur dengan baik.

5.3. Saran

Hasil penelitian ini menyarankan bahwa langkah lebih baik apabila penelitian di masa depan dapat menyelidiki pengembangan instrumen tes tertulis pada konsep IPA yang lain. Misalnya, penting untuk menentukan perbedaan hasil analisis butir soal yang satu dengan butir soal yang lain. Terkait dengan butir soal yang harus bermakna dengan menggunakan data yang *real* yang didapat dari *project* atau percobaan saintifik yang memang dilakukan oleh peneliti sebelum menyusun

instrumen tes ini. Hal ini dilakukan agar penyusunan instrumen soal yang disusun dapat menggalikan keterampilan *STEM* yang sesungguhnya diukur pada siswa. Penelitian di masa depan juga dapat membahas mengenai pengembangan instrumen tes untuk mengukur keterampilan *STEM* namun dari satu persatu keterampilan misalnya, sehingga akan dihasilkan penelitian yang lebih komprehensif. Untuk peneliti lainnya, disarankan untuk melakukan uji coba instrumen tes tertulis untuk mengukur keterampilan *STEM* dengan subjek yang lebih luas. Hal ini dimaksudkan agar tingkat keterbacaan, validitas dan reliabilitas dan keefektifan instrumen tes tersebut. Analisis model *Rasch* juga dapat dilakukan hingga analisis opsi jawaban, hal ini merupakan kelemahan yang tidak dilakukan oleh penelitian ini.

Studi saat ini hanya memperkirakan bahwa perbedaan kurikulum, praktik pembelajaran, dan kualitas guru adalah beberapa faktor yang berkontribusi pada item DIF (*Differential Item Functioning*) atau deteksi bias item. Apabila ingin meneliti lebih jauh mengenai deteksi bias item, barangkali penelitian selanjutnya harus mencoba untuk menyelidiki sumber DIF menggunakan data yang sama sehingga intervensi yang sesuai dapat dilakukan untuk meningkatkan kualitas desain pengujian. Prosedur dapat diperiksa lebih teliti untuk menentukan sumber potensial lain yang tidak ditemukan dalam penelitian ini. Analisis distraktor menggunakan *Rasch* juga perlu dilakukan dalam penelitian berikutnya agar informasi yang diperoleh berguna bagi penyempurnaan dan pengembangan soal menjadi lebih akurat.

