

BAB V

SIMPULAN, IMPLIKASI, DAN REKOMENDASI

5.1 Simpulan

Penelitian dan pengembangan telah dilakukan melalui tahapan pengkajian literatur dan informasi yang relevan, uji coba, dan implementasi dari produk yang dihasilkan. Instrumen berbentuk soal uraian yang dihasilkan memiliki autentisitas lebih baik daripada soal pilihan ganda dan disusun dengan konten yang kontekstual. Instrumen soal telah diuji secara kualitatif dan kuantitatif, dengan hasil uji coba menunjukkan bahwa kriteria valid dan reliabel telah terpenuhi. Instrumen asesmen yang dikembangkan juga mencakup rubrik penskoran jawaban yang telah diuji dan terbukti valid secara isi dan memiliki reliabilitas antar penskor dalam kategori tinggi. Penelitian juga mencakup kajian performa penskor jawaban uraian otomatis, yaitu UKARA, dengan hasil uji reliabilitas dalam kategori tinggi. Hasil asesmen dalam penelitian ini dapat efektif sebagai bagian dari evaluasi dan pengembangan program peningkatan kompetensi guru yang berbasis TPACK dalam skala luas. Simpulan hasil penelitian lebih rinci dijelaskan seperti berikut ini.

Pertama, instrumen soal memiliki autentisitas lebih tinggi secara konseptual, karena jawaban uraian dari soal yang diberikan dikonstruksi sendiri oleh masing-masing subjek berdasarkan pengetahuan yang dimilikinya. Instrumen soal memiliki konten yang kontekstual karena disusun berdasarkan materi IPA SMP sesuai dengan kurikulum yang berlaku dan tipe instrumen berupa soal-soal konseptual dan pemecahan masalah, dengan permasalahan yang terjadi dalam konteks pembelajaran IPA di SMP. Instrumen juga memenuhi karakter kontekstual karena disusun berdasarkan kerangka kerja TPACK yang sesuai dengan perkembangan pendidikan pada abad 21 sekarang ini. Uji validitas isi secara kualitatif dan uji validitas konstruk secara kuantitatif melalui prosedur EFA dan CFA telah dilakukan dalam penelitian dan hasilnya instrumen soal memenuhi kriteria valid. Reliabilitas instrumen soal berdasarkan koefisien Alpha Cronbach termasuk ke dalam reliabilitas tinggi. Kerangka kerja TPACK pada penelitian ini sesuai dengan pandangan transformatif dan dibangun oleh 4 aspek, yaitu PCK,

TCK, TPK, dan TPCK. Instrumen soal juga lebih efisien untuk mengukur tingkat TPACK dalam asesmen skala luas dibandingkan instrumen dalam bentuk lain sebelumnya yang lebih bersifat kualitatif sehingga kurang efisien.

Kedua, kualitas rubrik penskoran yang digunakan untuk menilai jawaban pada asesmen TPACK guru IPA tersusun secara sistematis oleh: (1) bagian penjelasan konsep yang mendasari permasalahan pada soal, rentang skor jawaban 0 – 2; (2) kriteria untuk masing-masing skor yang jelas dan rinci; dan (3) contoh jawaban yang dapat diberikan skor 2 sebagai skor tertinggi. Struktur rubrik tersebut terbukti valid, reliabel, dan efektif berdasarkan uji coba yang telah dilakukan. Instrumen rubrik telah melalui proses validasi oleh pakar, sama seperti untuk instrumen soal, sehingga memiliki validitas konten yang memadai. Rubrik penskoran yang telah disusun pada penelitian ini memiliki nilai reliabilitas antar penskor pada kategori tinggi untuk Intraclass Correlation Coefficient (ICC), sehingga telah terbukti memiliki reliabilitas yang memadai sebagai bagian dari instrumen asesmen pengetahuan guru. Efektifitas rubrik penskoran dibuktikan melalui penggunaan rubrik sebagai panduan oleh empat orang penskor dan menghasilkan skor yang kemudian layak untuk diuji lebih lanjut.

Ketiga, dari penelitian ini diperoleh hasil berupa gambaran performa penskor otomatis yang diuji coba untuk mendukung efisiensi sumber daya dalam penggunaan instrumen soal uraian terbatas. Penskoran jawaban uraian menggunakan penskor otomatis UKARA telah terbukti efektif dalam tahapan pelatihan untuk penelitian ini. Efektifitas penskoran berbantuan komputer tersebut dibuktikan melalui uji reliabilitas antar penskor, yaitu reliabilitas antara penskor manusia dan UKARA. Reliabilitas antar penskor yang tinggi tersebut meliputi ICC yang mencapai 0,77, koefisien Alpha (Cronbach) sebesar 0,8. Nilai reliabilitas antar penskor ICC tinggi menunjukkan bahwa UKARA dapat memberikan hasil penskoran yang relatif sama secara statistik dengan hasil penskoran yang dilakukan oleh manusia. Koefisien Alpha menjadi indikator konsistensi internal dalam penskoran yang dilakukan. Hasil implementasi juga menunjukkan performa yang relatif sama dengan hasil pelatihan atau uji coba penskor otomatis sebelumnya, yaitu mencapai nilai reliabilitas antar penskor yang tinggi dan persentase kesepakatan yang memadai. Penggunaan penskor otomatis

UKARA ini menjadi daya dukung untuk model instrumen yang telah dikembangkan karena menjadi bukti bahwa instrumen esai dapat memiliki karakter efisien dalam penskoran, seperti juga instrumen pilihan ganda.

Keempat, nilai yang ditunjukkan sebagai hasil asesmen TPACK guru IPA rata-rata berada pada kategori yang termasuk rendah. Hasil implementasi instrumen soal uraian dan penskoran berbantuan komputer untuk TPACK guru IPA SMP menunjukkan rerata nilai 34,55 untuk 250 orang subjek. Nilai untuk masing-masing aspek penyusun instrumen pengukuran, yaitu rerata nilai PCK 28,80, rerata nilai TCK 39,73, rerata nilai TPK 37,00, dan rerata nilai TPCK 32,67. Hasil penelitian tersebut tidak memiliki korelasi yang kuat dengan latar belakang pendidikan, masa kerja, usia, dan kelompok gender subjek berdasarkan hasil analisis korelasi (Pearson). Hasil tes berdasarkan latar belakang pendidikan secara berurutan dari yang tertinggi adalah kelompok guru dengan latar belakang pendidikan kurang sesuai, guru dengan latar belakang pendidikan sesuai, dan guru dengan latar belakang pendidikan tidak sesuai. Perbandingan hasil pengukuran berdasarkan masa kerja, nilai rerata tertinggi adalah guru dengan masa kerja lebih dari 10 tahun, berikutnya adalah untuk guru dengan masa kerja kurang dari 5 tahun, dan paling rendah kelompok guru dengan masa kerja antara 5 – 10 tahun. Untuk perbandingan hasil asesmen berdasarkan usia subjek, nilai rerata guru-guru pada kelompok dengan usai kurang dari 30 tahun mencapai nilai rerata paling tinggi, kemudian kelompok usia 30 – 45, dan guru-guru pada kelompok usia lebih dari 45 tahun. Nilai rerata TPACK untuk kelompok guru wanita lebih tinggi daripada nilai rerata guru laki-laki pada penelitian ini. Berdasarkan analisis dari hasil uji beda, nilai hasil tes untuk masing-masing variabilitas tersebut, yaitu latar belakang, masa kerja, usia, dan gender, tidak memiliki perbedaan secara statistik. Hasil implementasi ini menjadi bukti autentik bahwa tingkat pengetahuan TPACK guru IPA masih perlu ditingkatkan karena masih rata-rata nilai yang diperoleh masih pada kategori rendah.

5.2 Implikasi

Hasil penelitian ini terutama berimplikasi terhadap strategi asesmen yang dilakukan terhadap para guru. Instrumen berupa soal uraian terbatas dapat menjadi alternatif yang lebih baik daripada instrumen sebelumnya untuk mengukur pengetahuan guru IPA. Model instrumen yang dihasilkan relatif dapat menggantikan instrumen yang telah digunakan pada uji kompetensi guru atau asesmen lain terkait pengetahuan guru yang berbentuk pilihan ganda. Autentisitas yang lebih baik, kontekstual, dan tetap efisien menjadi karakteristik yang mendukung implikasi tersebut. Selain itu, secara umum model instrumen yang dikembangkan dalam penelitian ini dapat menjadi model untuk meningkatkan kualitas instrumen serta penggunaan instrumen alternatif yang lebih moderat.

Implikasi kedua dari hasil penelitian pengembangan ini adalah pemanfaatan teknologi penskoran otomatis dapat digunakan lebih luas lagi untuk meningkatkan efisiensi sumber daya sesuai dengan karakteristik pendidikan abad 21. Teknologi yang dimaksud terutama untuk sistem asesmen kompetensi guru. Penskoran esai berbantuan komputer berbasis kecerdasan buatan (AI) adalah salah bentuk pemanfaatan teknologi tersebut. Oleh karena itu penggunaan komputer seharusnya menjadi dukungan untuk penerapan instrumen asesmen alternatif dari asesmen yang lebih tradisional karena memiliki keterbatasan dalam autentisitas, konsistensi serta kurang efisien. Efisiensi ini menjadi faktor utama yang mendukung sistem evaluasi dalam skala luas atau nasional.

Hasil asesmen dalam penelitian ini mungkin tidak secara otomatis dapat mempengaruhi kebijakan dalam program evaluasi guru. Implikasi berdasarkan gambaran nilai pada penelitian ini dapat terkait dengan pengembangan keprofesian berkelanjutan para guru. Pengukuran kemampuan atau pengetahuan guru dengan kerangka kerja TPACK dan hasil yang lebih autentik dapat menjadi dasar yang lebih bermakna untuk program peningkatan kualitas guru. Mengetahui tingkat pengetahuan TPACK akan berimplikasi terhadap program pelatihan bagi guru-guru yang telah bertugas serta peningkatan kualitas kurikulum di perguruan tinggi bagi para calon guru, khususnya mengenai pemanfaatan teknologi dalam pembelajaran.

5.3 Rekomendasi

Hasil penelitian ini antara lain menunjukkan bahwa tingkat TPACK guru-guru SMP yang menjadi partisipan mendapat rerata nilai pada kategori rendah. Hasil tersebut dapat menjadi gambaran bahwa kemungkinan kondisi yang serupa tidak hanya terjadi pada daerah yang menjadi konteks dalam penelitian ini, yaitu Provinsi Banten, tetapi dapat pula terjadi pada lingkup nasional atau daerah lain. Pengetahuan pada kategori yang masih belum sesuai harapan tersebut juga telah diketahui terjadi pada asesmen kompetensi guru sebelumnya. Oleh karena itu pemerintah maupun pengelola pendidikan tinggi penyelenggara program atau pendidikan keguruan dapat menerapkan strategi asesmen yang lebih berkualitas dan moderat untuk mendapatkan informasi yang lebih akurat dan presisi sebagai dasar pengembangan program peningkatan kompetensi guru atau calon guru.

Gambaran yang diberikan pada hasil penelitian dan implementasi instrumen ini dapat ditindaklanjuti oleh pihak yang berwenang, yaitu pemerintah daerah dan komunitas guru IPA. Hasil penilaian TPACK dapat dijadikan sebagai dasar pengembangan program peningkatan kompetensi guru, khususnya untuk guru IPA di wilayah Provinsi Banten. Untuk komunitas guru IPA atau dikenal dengan Musyawarah Guru Mata Pelajaran (MGMP), hasil-hasil penelitian seperti ini dapat menjadi dasar untuk pengembangan kegiatan akademis dalam skala komunitas atau kelompok di masing-masing wilayah sehingga lebih terarah dan bermkna untuk peningkatan profesionalisme guru secara berkelanjutan.

Hasil penilaian ini juga dapat dijadikan salah satu dasar dalam evaluasi program pendidikan calon guru IPA di lembaga penyelenggara pendidikan calon guru. Data-data dasar yang lebih autentik dapat menjadi landasan perbaikan dalam program atau kurikulum pendidikan calon guru yang akan berdampak lebih jauh terhadap kualitas hasil pendidikan di Indonesia. Kerangka kerja TPACK adalah suatu konsep yang dianggap sesuai dengan perkembangan zaman, jadi seharusnya kurikulum pendidikan guru juga semakin berhubungan dengan pemanfaatan teknologi dalam pembelajaran.

Keterbatasan dalam proses penelitian dan pengembangan ini antara lain adalah jumlah partisipan, konteks wilayah partisipan, materi pelajaran yang tercakup dalam instrumen, dan keterbatasan dalam penggunaan penskoran

otomatis. Jika model instrumen yang telah dikembangkan dalam penelitian akan ditingkatkan terkait keterbatasan-keterbatasan tersebut, beberapa perbaikan dapat dilakukan dalam penelitian selanjutnya. Penelitian dan pengembangan yang akan dilakukan selanjutnya dapat mencakup beberapa kegiatan yang lebih baik lagi. Pertama, hasil penelitian ini dapat menjadi dasar pengembangan instrumen dan penggunaan sistem penskoran sejenis dalam skala yang lebih luas oleh pihak yang berwenang dalam sistem evaluasi kompetensi guru. Kedua, keterlibatan lebih banyak guru atau calon guru sebagai partisipan dalam penelitian pengembangan instrumen serupa untuk mata pelajaran yang berbeda perlu difasilitasi oleh pihak yang berkompeten, sehingga dapat meningkatkan keyakinan terhadap hasil uji secara statistik atau kuantitatif. Ketiga, guru-guru dalam satu wilayah kerja yang sama kemungkinan memiliki keseragaman dalam kemampuan dan penggunaan bahasa, jika penelitian dapat melibatkan guru-guru atau calon guru di suatu kampus pendidikan yang bersifat lintas wilayah, maka data penelitian akan semakin kaya dan instrumen hasil penelitian akan semakin kontekstual secara nasional karena adanya variabilitas tersebut. Keempat, instrumen yang telah disusun ini membatasi materi pelajaran hanya pada dua materi IPA sebagai bagiannya, jika instrumen yang dikembangkan, dalam mata pelajaran apapun, menjadikan materi pelajaran lebih banyak atau lebih bervariasi maka hasil yang diperoleh juga mungkin berbeda dan lebih baik. Kelima, penelitian untuk asesmen TPACK dapat dilakukan secara korelasional antara tingkat pengetahuan dan performa dalam praktik pembelajaran serta pengembangan sistem asesmen yang lebih komprehensif sesuai dengan karakteristik kerangka kerja TPACK. Keenam, perangkat lunak UKARA yang digunakan dalam penelitian ini tidak sepenuhnya dikembangkan untuk menjadi penskor dalam asesmen guru, oleh karena itu pengembangan sistem komputer untuk penskor otomatis yang khusus untuk asesmen kompetensi guru yang lebih kompleks akan memiliki potensi untuk meningkatkan kualitas hasil penelitian yang dilakukan.