Available online at: http://journal.uny.ac.id/index.php/jpms

**Jurnal Pendidikan Matematika dan Sains, 8 (2), 2020, 93-102**

# Pengembangan Instrumen *Integrated Assessment* antara Keterampilan Proses Sains dan Penguasaan Konsep IPA untuk Mengukur Aspek Kognitif Proses

**Agus Santoso*, Anti Kolonial Prodjosantoso**

Pendidikan Sains, Program Pascasarjana, Universitas Negeri Yogyakarta. Jalan Colombo Yogyakarta No.1, Karang Malang, Sleman, Daerah Istimewa Yogyakarta 55281, Indonesia

*Korespondensi Penulis. E-mail: agus4sains@gmail.com; prodjosantoso@uny.ac.id

**Abstrak**

Penelitian ini bertujuan untuk mendeskripsikan kelakayan instrumen, mengetahui hasil pengukuran aspek kognitif proses, dan mengetahui respon keterpakaian instrumen. Penelitian ini merupakan *research & development* dengan menggunakan model pengembangan 4D. Prosedur penelitian ini meliputi (1) *define*, (2) *design*, (3) *develop*, dan (4) *disseminate*. Subjek uji coba validasi *instrumen integrated assessment* dilakukan pada siswa di satu SMP dan uji keterpakaian produk pada siswa di tiga SMP di provinsi DIY. Teknik pengambilan data meliputi wawancara, angket, dan tes. Hasil penelitian ini berupa produk soal pilihan ganda untuk mengukur aspek kognitif proses yang layak secara teoritis ditinjau dari standar pembuatan tes sesuai hasil validasi ahli dan uji coba empiris. Hasil validasi ahli diperoleh validasi isi V Aiken sebesar 0,98 yang berarti instrumen tes dinyatakan sangat baik, sedangkan hasil uji coba empiris diperoleh 49 butir soal fit dan satu butir soal tidak fit, selanjutnya reliabilitas item tes adalah 0,73 dan reliabilitas sampel adalah 0,92. Hasil pengukuran di sekolah menunjukkan sekolah berkategori tinggi memiliki penguasaan aspek kognitif proses lebih baik daripada sekolah berkategori sedang dan kategori rendah. Hasil respon keterpakaian produk instrumen *integrated assessment* di sekolah dinyatakan dengan kriteria baik.

**Kata Kunci**: instrumen *integrated assessment*, keterampilan proses sains, penguasaan konsep IPA, aspek kognitif proses

# *Developing Instrument of Integrated Assessment between Science Process Skills and Science Mastery Concepts for Measuring Aspects of Process Cognitive*

***Abstract***

*This study aims to describe the feasibility of the instrument, determine the results of the measurement of aspects of processes cognitive, and study the response of the instrument utilization. This study is a research & development using 4D development model. The procedure of this study include (1) define, (2) design, (3) develop, and (4) disseminate. The test subject of validation instrument integrated assessment was performed on students at junior high school and test utilization of the product on three junior high schools in the province DIY. The techniques of collection data include interviews, questionnaires and tests. The results of this study are the product of multiple choice questions to measure aspects of processes cognitive theoretically feasible in terms of standard of making the test according to the results of expert validation and empirical testing. The results of expert validation obtained validation of the contents of Aiken's V 0.98 which means that the test instrument is otherwise very good, while the trial results empirically obtained 49 items fit and one items does not fit, then the reliability of the test items is 0,73 and the reliability of the test sample is 0,92. The results of measurements in schools show that in high school category tend to have mastery of the aspects of processes cognitive better than the schools categorized medium and low categories. The response results of product utilization instrument of integrated assessment in the school otherwise the result is in good criteria.*

***Keywords:*** *integrated assessment instrument, science process skills, science mastery concepts, processes cognitive aspect*

**How to Cite**: Santoso, A., & Prodjosantoso, A. K. (2020). Pengembangan instrumen *integrated assessment* antara keterampilan proses sains dan penguasaan konsep IPA untuk mengukur aspek kognitif proses. *Jurnal Pendidikan Matematika dan Sains, 8*(2), 93-102. doi:http://dx.doi.org/10.21831/jpms.v8i2.21225

***Permalink/DOI***: ***DOI:*** *http://dx.doi.org/10.21831/jpms.v8i2.21225*

Copyright © 2020, JPMS, p-ISSN: 1410-1866, e-ISSN: 2549-1458

## PENDAHULUAN

Pendidikan merupakan salah satu aspek yang menunjang perkembangan suatu negara dan merupakan proses sistematis dalam upaya meningkatkan martabat manusia secara holistis. Dalam hal ini, pendidikan merupakan sebuah aktualisasi dimensi kemanusiaan yang terdiri dari aspek kognitif, psikomotor, dan afektif (Salsabila & Firdaus, 2018). Aspek kognitif tercermin pada kapasitas pikir dan daya intelektualitas untuk menggali ilmu pengetahuan, mengembangkan, serta menguasai teknologi. Aspek psikomotor tercermin pada kemampuan mengembangkan keterampilan teknis dan kecakapan praktis. Aspek afektif tercermin pada kualitas keimanan dan ketakwaan, etika, estetika, akhlak mulia, dan budi pekerti luhur. Tujuan pendidikan seperti ini dapat tercapai dengan proses pembelajaran yang baik.

Proses pembelajaran terdiri dari tiga tahap, yaitu perencanaan, pelaksanaan proses pembelajaran, dan penilaian hasil pembelajaran. Salah satu indikator untuk melihat keberhasilan pencapaian tujuan pembelajaran adalah hasil penilaian. Hasil penilaian mengungkapkan informasi lengkap dan sesuai dengan data yang diperlukan. Sementara itu, hasil penilaian yang lengkap dan sesuai dapat diperoleh dengan menggunakan instrumen penilaian yang tepat. Standar penilaian pendidikan di Indonesia diatur dalam Permendikbud Nomor 66 Tahun 2013, yaitu penilaian tidak hanya didasarkan pada kompetensi dasar (KD), namun juga pada standar kompetensi lulusan (SKL) dan kompetensi inti (KI). Hal tersebut sesuai dengan yang tertulis dalam pedoman Kementerian Pendidikan dan Kebudayaan (2013) yang juga menyatakan bahwa penilaian tidak hanya pada level kompetensi dasar, tetapi juga berdasarkan KI dan SKL. Perubahan dalam standar penilaian pendidikan tersebut akan mempengaruhi proses penilaian yang dilakukan guru, salah satunya guru mata pelajaran IPA di SMP.

Lebih lanjut, Trianto (2014) menyatakan bahwa IPA dibangun atas dasar produk ilmiah, proses ilmiah, dan sikap ilmiah. Hal yang serupa juga dicetuskan oleh Kuswanto (2008) bahwa pada dasarnya tujuan IPA adalah menguasai pengetahuan IPA, memahami dan menerapkan konsep IPA, menerapkan keterampilan proses, dan mengembangkan sikap. Sementara itu, tujuan pendidikan IPA dalam Permendikbud no. 68 tahun 2013 yaitu menekankan pemahaman lingkungan dan alam sekitar beserta kekayaan yang dimilikinya yang perlu dilestarikan dan dijaga dalam perspektif biologi, fisika, dan kimia. Dari hal tersebut dapat dikatakan bahwa tujuan pendidikan IPA adalah menekankan pemahaman konsep IPA dan penerapan keterampilan proses dalam pelestarian lingkungan dan alam sekitar pada perspektif biologi, fisika, dan kimia. Dengan demikian, pembelajaran IPA mencakup konten dari biologi, fisika, dan kimia yang diajarkan dengan cara terpadu.

Pembelajaran IPA membutuhkan instrumen penilaian yang sesuai. Sementara itu, Arifin (2013) menyatakan bahwa salah satu karakteristik instrumen yang baik adalah bersifat relevan. Relevan berarti instrumen yang digunakan haruslah sesuai dengan standar kompetensi, kompetensi dasar, dan indikator yang telah ditetapkan. Dalam konteks penilaian hasil belajar, instrumen harus sesuai dengan domain hasil belajar yang meliputi domain kognitif, afektif, dan psikomotor (Subagia & Wiratma, 2016). Oleh sebab itu, materi yang diajarkan dengan konteks penilaian hasil belajar haruslah relevan dengan bentuk penilaian dan domain hasil belajar yang dinilai guru sesuai dengan hakikat IPA. Lebih lanjut, Sardinah et al. (2012) menyatakan bahwa IPA pada hakikatnya merupakan produk, proses, sikap, dan aplikasi. Sebagai produk, IPA merupakan kumpulan pengetahuan dan konsep. Sebagai proses, IPA merupakan proses yang digunakan untuk mempelajari objek studi, menemukan, dan mengembangkan produk sains.

Sebagai sikap, IPA merupakan cara pandang dan sikap dalam ilmu pengetahuan. Sebagai aplikasi, IPA akan melahirkan teknologi yang dapat memberikan kemudahan bagi kehidupan. Oleh sebab itu, penilaian seyogyanya mengacu pada hakikat IPA yaitu menilai IPA sebagai produk, proses, sikap, dan aplikasi. Akan tetapi, hasil pengamatan di lapangan menunjukkan penilaian terhadap keterampilan proses jarang dilakukan oleh guru IPA. Sulitnya guru memanajemen waktu dan banyaknya siswa yang akan dinilai menyebabkan penilaian keterampilan proses jarang dilakukan. Untuk menilai keterampilan proses dapat digunakan cara nontes dengan menggunakan lembar pengamatan (Usman, 2006). Namun demikian, penggunaan lembar pengamatan berdasarkan hasil wawancara dirasa memberatkan guru.

Copyright © 2020, JPMS, p-ISSN: 1410-1866, e-ISSN: 2549-1458

Lebih lanjut, Arini et al. (2017) menyatakan bahwa agar tidak memberatkan guru, pelaksanaan penilaian keterampilan proses dapat dilakukan secara bertahap. Hal ini terjadi karena tidak mungkin satu guru dapat mengamati banyak siswa dalam waktu yang bersamaan. Cara penilaian seluruh siswa dapat dilakukan, apabila ada observer yang membantu melakukan pengamatan selama penilaian. Akan tetapi, keadaan di sekolah tidak memungkinkan adanya observer karena guru IPA yang lain terkadang sibuk dengan tugas mengajar. Kegiatan penilaian yang dilakukan guru di sekolah dapat dilakukan dengan berbagai cara, salah satunya dalam bentuk *paper and pencil test.* Kondisi di lapangan menujukkan bahwa selama ini, *paper and pencil test* cenderung mengukur kemampuan kognitif tanpa memperhatikan pada penilaian keterampilan proses sains siswa. Padahal seharusnya penilaian keterampilan proses sains juga penting dilakukan.

Hal ini akan mengubah paradigma guru dalam penilaian terhadap siswa, yaitu pada penilaian bentuk *paper and pencil test* yang cenderung mengukur kemampuan kognitif dan jarang memperhatikan kemampuan keterampilan proses, menuju pada penilaian yang sesuai dengan hakikat IPA. Penilaian IPA dalam bentuk tes standar international, salah satunya terwujud dalam bentuk tes yang dilakukan oleh *Program for International Student Assessment (PISA)* dan *The Learning Curve Pearson 2014.* Tujuan tes tersebut adalah menentukan patokan untuk melihat dan menentukan seberapa jauh siswa dapat bersaing dalam era globalisasi. Tes ini dalam pembelajaran sains biasanya mengukur kemampuan berpikir tingkat tinggi dan keterampilan proses sains siswa yang terwujud dalam literasi sains. Sering kali hasil tes tersebut digunakan untuk mendasari kajian dalam melakukan program pengembangan dan peningkatan mutu pendidikan di beberapa negara.

Berdasarkan hasil PISA tahun 2012, Indonesia mendapatkan peringkat 64 dari 65 negara anggota PISA dengan skor literasi sains sebesar 382 (Winata & RW, 2018). Berdasarkan hasil PISA tersebut dapat diketahui bahwa literasi sains di Indonesia mendapatkan nilai yang kurang bagus. Berdasarkan hasil *The Learning Curve Pearson 2014* yang menggambarkan indeks global kemampuan kognitif dan hasil pendidikan, Indonesia berada pada posisi terendah yaitu peringkat 40 dari keseluruhan negara anggotanya. Hal ini menggambarkan bahwa kualitas pendidikan di Indonesia masih rendah dibandingkan dengan negara lainnya. Salah satu indikator untuk melihat keberhasilan pencapaian tujuan pembelajaran adalah hasil penilaian. Hasil penilaian haruslah mengungkapkan informasi lengkap dan sesuai dengan data yang diperlukan. Untuk mendapatkan hasil penilaian yang tepat seharusnya digunakan instrumen penilaian yang tepat sesuai dengan pembelajaran.

Pembelajaran IPA tidak bisa lepas dari praktikum. Praktikum dapat mambantu siswa untuk lebih memahami hakikat sains. Sementara itu, Sukarno dan Hamidah (2013) menyatakan bahwa siswa yang mendapatkan skor tes keterampilan proses rendah secara relatif jarang melakukan aktivitas pembelajaran yang berbasis laboratorium. Hasil studi awal yang dilakukan dengan wawancara terhadap Guru IPA mengenai pembelajaran IPA di beberapa SMP di Daerah Istimewa Yogyakarta, menunjukkan bahwa sebagian besar guru merasa kesulitan dalam mengukur keterampilan proses sains pada siswa karena terkendala keterbatasan waktu dan jumlah siswa yang banyak. Selain itu, guru juga harus membimbing praktikum sekaligus menilai kemampuan siswa dalam kelas dalam waktu yang bersamaan. Oleh karena itu, guru di lapangan menilai siswa secara subjektif dengan melakukan estimasi tanpa menggunakan instrumen yang valid, yaitu dengan melihat siswa yang aktif dan pasif.

Siswa yang aktif diberi nilai tinggi dan yang pasif diberi nilai rendah, sedangkan yang antara pasif dan aktif diberi nilai sedang. Hal ini menyebabkan penilaian tidak sesuai karena bersifat subjektif. Penilaian yang bersifat subjektif tidak sesuai dengan tuntutan pedoman kurikulum. Penilaian seharusnya dilakukan dengan instrumen yang jelas dan dilakukan secara objektif. Oleh karena itu, masalah sulitnya melakukan penilaian harus dicari solusinya, salah satunya dengan mengembangkan instrumen penilaian terintegrasi. Bentuk penilaian terintegrasi antara keterampilan proses sains dan penguasaan konsep IPA dipilih karena ada hubungan antar keduanya. Selain itu, kegiatan penilaian dilakukan setelah proses pembelajaran berlangsung, sehingga dalam diri siswa sudah belajar dari segi konsep dan keterampilan.

Copyright © 2020, JPMS, p-ISSN: 1410-1866, e-ISSN: 2549-1458

Berdasarkan hasil wawancara, sebagian besar guru menyatakan bahwa masih belum banyak ditemui contoh soal IPA yang secara terintegrasi mengukur kemampuan kognitif dan keterampilan proses sains siswa. Hasil wawancara tersebut mengindikasikan perlu adanya bentuk instrumen penilaian IPA yang dapat mengukur keterampilan proses sains dan penguasaan konsep siswa. Lebih lanjut, Darmayanti et al. (2013) menyatakan bahwa terdapat indikator keterampilan proses dalam domain kognitif. Bentuk instrumen penilaian perlu dianalisis agar dapat mengevaluasi pembelajaran dan kemampuan siswa. Pertanyaan dalam soal juga digunakan untuk mengevaluasi pembelajaran yang telah dilaksanakan. Penilaian pembelajaran IPA yang memuat domain kognitif (penguasaan konsep IPA) dan keterampilan proses sains dalam pembelajarannya, memerlukan bentuk soal yang relevan dan mengarah pada penilaian kognitif dan keterampilan proses sains.

Sementara itu, Suryandari (2013) menyatakan bahwa penilaian terhadap keterampilan proses dapat dilakukan dengan cara tes tertulis, meskipun tes ini tidak dapat menjangkau semua kemampuan karena penggunaan indera pendengaran dan peraba tidak mungkin dinilai dengan tes tertulis. Instrumen untuk mengukur keterampilan proses sains dapat dilakukan dengan soal yang berbentuk pilihan ganda. Selain itu, Saputro et al. (2019) telah mengembangkan *integrated assessment* antara keterampilan proses sains dan penguasaan materi ajar listrik dan magnet. Hal serupa juga telah dilakukan oleh Shahali dan Halim (2010) yang telah mengembangkan instrumen tes pilihan ganda untuk mengukur keterampilan proses sains.

Sejalan dengan hal tersebut, Leighton dan Gierl (2011) menyatakan banyak penelitian pembelajaran yang dilakukan pada pengetahuan ilmiah dan keterampilan telah difokuskan pada perolehan keterampilan proses. Jadi, merupakan hal baru dalam penelitian pengembangan pembelajaran yang difokuskan pada perolehan keterampilan proses sains sekaligus penguasaan konsep. Pengembangan instrumen penilaian ini untuk memberikan solusi terhadap permasalahan yang telah diungkapkan, yaitu dengan mengembangkan bentuk penilaian terintegrasi karena adanya hubungan antara keterampilan proses sains dan penguasaan konsep IPA.

Penjelasan tersebut mendukung pengembangan bentuk instrumen penilaian IPA yang disebut dengan instrumen *integrated assessment.* Penelitian pengembangan instrumen *integrated assessment* antara keterampilan proses sains dan penguasaan konsep IPA berusaha memberikan kontribusi dalam bentuk penilaian pembelajaran IPA yang dapat mengukur indikator keterampilan proses sains sekaligus indikator penguasaan konsep IPA yang termasuk ke dalam penilaian untuk mengukur aspek kognitif proses siswa SMP. Hasil pengukuran tes ini diharapkan diperoleh potret kemampuan aspek kognitif proses pada siswa SMP yang dapat digunakan untuk keperluan kebijakan lebih lanjut.

## METODE

Penelitian ini merupakan penelitian pengembangan atau *Research and Development* (R&D). Model pengembangan yang digunakan dalam penelitian ini adalah model pengembangan 4-D yang dikemukakan oleh Thiagarajan et al. (1974). Prosedur penelitian pengembangan ini meliputi tahap *define, design, develop,* dan *disseminate.* Sementara itu, penelitian ini dilaksanakan antara bulan Januari sampai Mei 2015 yang bertempat di SMPN 3 Sentolo, SMPN 1 Sleman, SMPN 3 Depok Sleman, dan SMPN 3 Tempel Sleman. Subjek penelitian saat uji coba lapangan pendahuluan adalah siswa kelas VII B, VII C, dan VII F SMPN 3 Sentolo. Subjek ketika uji coba keterpakaian produk adalah siswa kelas VII di SMPN 1 Sleman, SMPN 3 Depok Sleman, dan SMPN 3 Tempel Sleman. Setiap sekolah diambil masing-masing dua kelas untuk sampel penelitian.

Teknik pengumpulan data berupa tes dan non tes. Teknik non tes melalui wawancara dan angket, sedangkan teknik tes menggunakan soal pilihan ganda. Pengumpulan data dilakukan menggunakan instrumen angket, wawancara, soal pilihan ganda, dan kisi-kisi instrumen penilaian. Lebih lanjut, draf awal yang divalidasi menghasilkan data kualitatif. Data tersebut berupa saran perbaikan pada tiap butir soal yang berasal dari validator. Hasil penilaian validator dianalisis menggunakan rumus Aiken'V untuk menentukan validasi isi. Saran perbaikan dari validator dipadukan dan digunakan untuk merevisi draf awal. Selanjutnya, draf II yang di validasi menghasilkan data kualitatif.

Copyright © 2020, JPMS, p-ISSN: 1410-1866, e-ISSN: 2549-1458

Data tersebut berupa saran perbaikan pada setiap butir soal yang berasal dari praktisi. Saran perbaikan dari praktisi dipadukan dan digunakan untuk merevisi draf II yang sudah diberikan. Data kualitatif ketiga adalah hasil respon guru terhadap produk instrumen *integrated assessment*. Analisis kuantitatif untuk memberikan gambaran tentang validitas empiris merupakan penelaahan butir soal berdasarkan pada karakteristik internal tes melalui data yang diperoleh secara empiris. Karakteristik internal yang dimaksud meliputi kriteria *fit* dengan model *Rasch*, melihat nilai *infit t*, validitas soal, reliabilitas soal, tingkat kesukaran, daya beda, dan analisis pola persebaran jawaban. Analisis menggunakan teori klasik dan modern dengan bantuan program Quest. Hasil analisis menggunakan teori klasik dalam penelitian ini diperoleh informasi mengenai tingkat kesukaran, daya beda, dan analisis pola persebaran jawaban. Hasil analisis menggunakan teori modern diperoleh informasi mengenai kriteria *fit* dengan model *Rasch*, nilai *infit t*, serta reliabilitas item tes dan reliabilitas sampel.

## HASIL DAN PEMBAHASAN

Hasil validasi didapatkan melalui instrumen validasi dan butir soal yang telah dibuat. *Judgement* dari validasi ahli dibagi menjadi tiga, yaitu valid tanpa revisi, valid dengan revisi, dan tidak valid. Kriteria saran ketika terjadi revisi dibagi menjadi dua yaitu kriteria saran dari ahli evaluasi dan ahli materi. Hasil validasi dosen ahli evaluasi yang menyatakan bahwa 49 soal valid tanpa revisi dan 1 soal valid dengan revisi yaitu pada nomor 45. Selain itu, tidak ada soal yang tidak valid, sedangkan berdasarkan hasil validasi dosen ahli materi menyatakan 47 soal valid tanpa revisi dan 3 soal valid dengan revisi yaitu pada nomor 18, 46, dan 48, serta tidak ada soal yang tidak valid. Setelah saran perbaikan diperoleh, tahapan selanjutnya adalah dilakukan revisi terhadap draf awal instrumen *integrated assessment*. Draf awal diperbaiki sesuai dengan saran perbaikan yang telah diberikan dosen ahli.

Praktisi (Guru IPA) memberikan saran perbaikan pada draf I instrumen *integrated assessment*. Guru (praktisi pendidikan) memberikan masukan dari segi materi. Hasil validasi oleh guru pertama menyatakan bahwa 47 soal valid tanpa revisi dan 3 soal valid dengan revisi yaitu pada nomor 18, 26 dan 44, serta tidak ada soal yang tidak valid. Hasil validasi oleh guru kedua menyatakan bahwa 49 soal valid tanpa revisi dan 1 soal valid dengan revisi yaitu pada nomor 35, serta tidak ada soal yang tidak valid. Setelah saran perbaikan diperoleh, tahapan selanjutnya melakukan revisi terhadap draf I instrumen *integrated assessment*. Perbaikan draf I dilakukan sesuai saran perbaikan yang diberikan oleh praktisi. Penilaian dari keempat validator yaitu dua dosen ahli dan dua praktisi pendidikan dianalisis menggunakan rumus Aiken'V. Hasil analisis validitas isi terendah diperoleh sebesar 0,75 pada butir soal nomor 18, dan validitas isi tertinggi sebesar 1,00. Rata-rata validitas isi untuk semua butir soal didapatkan nilai 0,98 yang berarti instrumen *integrated assessment* yang dikembangkan sangat baik.

Uji coba instrumen merupakan tahapan validasi draf II instrumen *integrated assessment* secara empiris. Uji coba terbatas dilakukan dengan cara memberikan paket soal yang berbeda untuk menghindari contekan antar siswa. Paket soal A dan B secara substansi sama, namun urutan soal yang diacak antara paket soal A dan B. Analisis dengan program Quest menghasilkan nilai *reliability of item estimate* sebesar 0,92 yang artinya reliabilitas sampel sangat tinggi. Hasil *output* program Questnya dapat dilihat seperti pada Gambar 1.

```
Summary of Item Estimates
=========================

Mean                              .00
SD                                .84
SD (adjusted)                     .80
Reliability of estimate           .92


 Fit Statistics
================

 Infit Mean Square          Outfit Mean Square

    Mean     1.00             Mean     1.00
    SD        .07             SD        .10

      Infit t                    Outfit t

    Mean     -.01             Mean     -.01
    SD        .95             SD        .65
```

**Gambar 1.** Hasil *summary of item estimate*

Berdasarkan Gambar 1 dapat diketahui bahwa nilai reliabilitas sampel sangat tinggi karena nilainya 0,92. Hal tersebut menunjukkan bahwa semakin tinggi nilai reliabilitas sampel, maka semakin meyakinkan sampel uji coba sesuai atau cocok atau fit dengan item yang diujikan (Subali & Suyata, 2012). Hal ini berarti sampel untuk uji coba semakin baik untuk memberikan informasi yang diharapkan. Berdasarkan informasi untuk *fit statistics* diketahui bahwa besarnya *infit Mean Square (infit MNSC)* sebesar 1,0 dengan simpangan baku 0,07 artinya secara keseluruhan item sesuai dengan model *Rasch*. Sementara itu,

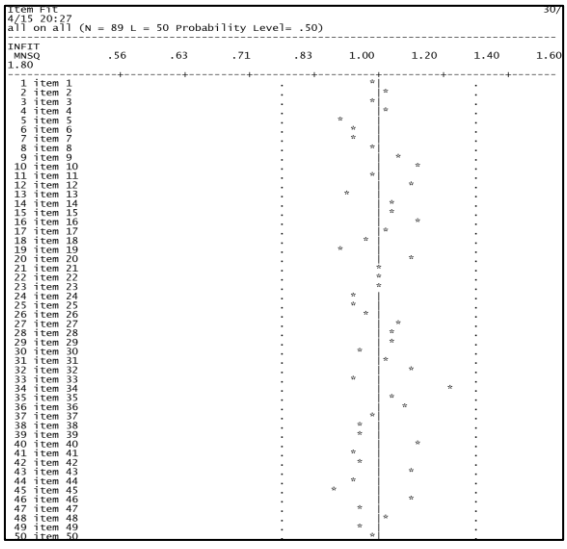
Copyright © 2020, JPMS, p-ISSN: 1410-1866, e-ISSN: 2549-1458

analisis dengan program Quest menghasilkan nilai *reliability of case estimate* sebesar 0,73 yang artinya reliabilitas tes tinggi untuk tes buatan guru. Hasil *output* program Questnya dapat dilihat pada Gambar 2 berikut.

```
Summary of case Estimates
=========================

Mean                              -.05
SD                                 .60
SD (adjusted)                      .52
Reliability of estimate            .73


 Fit Statistics
===============

 Infit Mean Square          Outfit Mean Square

    Mean     1.00              Mean     1.00
    SD        .13              SD        .18

      Infit t                    Outfit t

    Mean      .01              Mean      .04
    SD       1.12              SD        .81
```

**Gambar 2.** Hasil *summary of case estimate*

Gambar 2 memberikan informasi bahwa nilai reliabilitasnya sebesar 0,73. Hal ini berarti bahwa pengukuran memberikan hasil yang konsisten. Hal tersebut sesuai dengan pernyataan Surapranata (2005) bahwa koefisien reliabilitas 0,7 sampai 0,8 cukup tinggi untuk suatu penelitian dasar. Nilai reliabilitas tes akan rendah, jika pada suatu butir tertentu ada banyak siswa berkemampuan lebih tinggi justru salah mengerjakan, sebaliknya banyak siswa yang berkemampuan rendah yang benar mengerjakan (Sumintono & Widhiarso, 2015). Lebih lanjut, hasil nalisis lainnya diperoleh *mean infit* MNSQ sebesar 1,0 dan simpangan baku sebesar 0,13. Hal ini menunjukkan bahwa secara keseluruhan testi sesuai atau cocok atau *fit* dengan model *Rasch*. Informasi mengenai butir soal yang *fit* dengan model *Rasch* dapat disajikan pada Gambar 3.




**Gambar 3.** Butir soal *fit* dengan model *Rasch*

Pada Gambar 3 terlihat bahwa 50 butir soal *fit* atau cocok dengan model Rasch atau model 1-PL (*parameter logic*) dengan kriteria batas penerimaan ≥ 0,77 sampai ≤ 1,30. Langkah selanjutnya adalah melihat nilai *infit* t butir dengan kriteria -2,00 ≤ *t* ≤ 2,00 soal dinyatakan lolos, sedangkan jika nilai *infit t* diperoleh -2,00 > t > 2,00 maka soal dinyatakan gugur. Sementara itu, untuk butir soal nomor 34 mendekati garis batas, maka harus diperhatikan. Informasi mengenai nilai *infit* t soal nomor 34 disajikan pada Gambar 4 berikut.

| | ITEM NAME | SCORE | MAXSCR | THRSH 1 | INFT MNSQ | OUTFT MNSQ | INFT t | OUTFT t |
|---|---|---|---|---|---|---|---|---|
| 32 | item 32 | 29 | 89 | .72 .23 | 1.10 | 1.15 | 1.1 | 1.0 |
| 33 | item 33 | 53 | 89 | -.46 .22 | .93 | .90 | -1.0 | -.7 |
| 34 | item 34 | 34 | 89 | .46 .23 | 1.21 | 1.27 | 2.8 | 1.8 |
| 35 | item 35 | 34 | 89 | .46 .23 | 1.05 | 1.05 | .7 | .4 |

**Gambar 4.** Hasil *infit* t tiap butir soal

Pada Gambar 4 tampak pada butir soal 34 diperoleh *infit t* > 2,80 yang berarti soal dinyatakan gugur. Hal ini dapat terjadi karena besarnya *infit t* yang lebih dari 2,0 berarti soal ditolak atau tidak *fit* (Hasyim, 2015). Setelah dilakukan analisis menurut *IRT*, tahap selanjutnya adalah dilakukan analisis menggunakan metode *CTT*. Analisis menggunakan metode *CTT* dapat memberikan informasi tingkat kesukaran, daya beda dan pola persebaran jawaban, dan reliabilitas tes. Hasil analisis tingkat kesukaran menunjukkan bahwa draf instrumen *integrated assessment* memuat soal dalam kategori mudah sebanyak 8 soal (16%), kategori sedang sebanyak 36 soal (72%), dan kategori mudah sebanyak 6 soal (12%). Soal lebih banyak dengan kategori sedang karena ranah dalam pembuatan soal ini lebih banyak pada C2 hingga C4.

Hasil analisis daya beda menunjukkan bahwa instrumen *integrated assessment* sebanyak 58% berfungsi dengan baik, artinya dapat membedakan kemampuan siswa yang berkemampuan tinggi dengan berkemampuan rendah. Selanjutnya, 40% butir soal diperbaiki dari segi karakteristik soal dan pilihan jawaban. Lebih lanjut, untuk 1 soal yaitu nomor 34 memiliki daya pembeda yang negatif, sehingga siswa yang berkategori tinggi cenderung menjawab salah dan siswa yang berkategori rendah cenderung menjawab benar. Oleh sebab itu, butir soal nomor 34 dapat dihilangkan.

Copyright © 2020, JPMS, p-ISSN: 1410-1866, e-ISSN: 2549-1458

Hasil analisis pengecoh menunjukkan sebanyak 46 soal memiliki pemilih lebih dari 2,5% untuk setiap pilihan jawaban dengan persentase 92%, sedangkan untuk pemilih dibawah dari 2,5% untuk setiap pilihan jawaban dengan persentase 8%. Ketentuan tersebut sesuai Zulaiha (2010) yang menyatakan bahwa pengecoh dikatakan berfungsi apabila pengecoh dipilih paling sedikit oleh 2,5 % (≥ 0,025). Berdasarkan hasil analisis pengecoh dapat diketahui bahwa sebagian besar butir soal memiliki pengecoh yang berfungsi dengan baik. Sementara itu, rekapitulasi karakteristik butir soal dapat ditunjukkan pada Tabel 1 berikut.

**Tabel 1.** Rekapitulasi karakteristik butir soal

| No. | Kategori | Nomor Butir | Total | % |
|---|---|---|---|---|
| 1. | Diterima | 2, 3, 5, 6, 7, 8, 13, 18, 19, 22, 24, 25, 26, 30, 31, 33, 37, 38, 39, 41, 42, 44, 45, 47, 49, 50 | 26 | 52 |
| 2. | Diperbaiki | 1, 4, 9, 10, 11, 12, 14, 15, 16, 17, 20, 21, 23, 27, 28, 29, 32, 35, 36, 40, 43, 46, 48 | 23 | 46 |
| 3. | Ditolak | 34 | 1 | 2 |

Karakteristik butir soal pada Tabel 1 memberikan estimasi jumlah butir soal yang diterima sebanyak 26 butir, diperbaiki sebanyak 23 butir dan yang ditolak sebanyak 1 butir. Soal dengan kategori "diterima" langsung digunakan, soal dengan kategori "ditolak" tidak digunakan dan soal dengan kategori "diperbaiki" direvisi terlebih dahulu sebelum digunakan untuk pengukuran aspek kognitif proses pada siswa. Hasil uji coba keterpakaian instrumen *integrated assessment* di lapangan digunakan untuk mengukur kemampuan siswa dalam menguasai keterampilan proses sains sekaligus penguasaan konsep yang terintegrasi menjadi aspek kognitif proses.

Instrumen ini juga digunakan untuk memetakan aspek kognitif proses tersebut berdasarkan kategori sekolah rendah, sedang, dan tinggi. Jumlah butir soal dari produk instrumen *integrated assessment* setelah dilakukan validitas empiris menghasilkan 49 butir soal, dengan komposisi seperti pada Tabel 2 berikut.

**Tabel 2.** Instrumen *integrated assessment*

| No. | Aspek Kognitif Proses | Jumlah Butir |
|---|---|---|
| 1. | Mengobservasi | 4 |
| 2. | Kemampuan kuantifikasi | 5 |
| 3. | Menginferensi | 5 |
| 4. | Memprediksi | 5 |
| 5. | Mengklasifikasikan | 5 |
| 6. | Mengkomunikasikan hasil | 5 |
| 7. | Menganalisis hipotesis | 5 |
| 8. | Merencanakan percobaan | 5 |
| 9. | Menjelaskan penggunaan alat dan bahan | 5 |
| 10. | Mengidentifikasi variabel | 5 |

Berdasarkan hasil pengukuran, semua aspek kognitif proses siswa digabung untuk melihat potret aspek kognitif proses pada beberapa sekolah dengan perbedaan kategori sekolah. Perbandingan aspek kognitif proses antara tiga kategori sekolah yaitu kategori rendah, sedang, dan tinggi secara ringkas dapat dilihat pada Tabel 3 berikut.

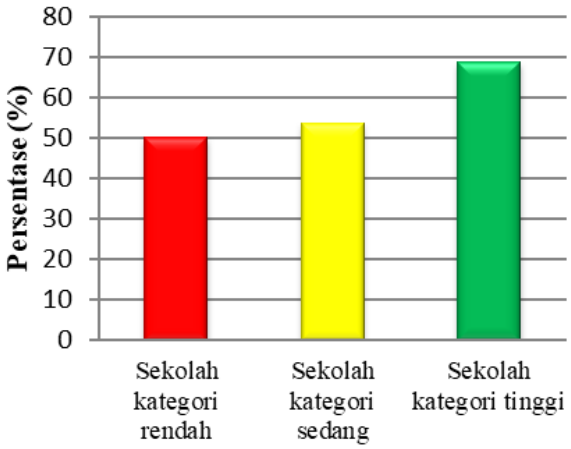
**Tabel 3.** Rerata penguasaan aspek kognitif proses

| No | Aspek Kognitif Proses | Persentase (%) | | |
|---|---|---|---|---|
| | | Sekolah Kategori Rendah | Sekolah Kategori Sedang | Sekolah Kategori Tinggi |
| 1 | Mengobservasi | 53,6 | 58,7 | 68,7 |
| 2 | Mengkuantifikasi | 57,7 | 54,3 | 69,8 |
| 3 | Menginferensi | 50,6 | 61,0 | 71,1 |
| 4 | Memprediksi | 42,9 | 44,4 | 47,9 |
| 5 | Mengklasifikasikan | 43,5 | 44,1 | 70,2 |
| 6 | Mengkomunikasikan Hasil | 60,3 | 62,9 | 84,4 |
| 7 | Menganalisis Hipotesis | 51,6 | 61,6 | 74,6 |
| 8 | Merencanakan Eksperimen | 41,9 | 38,7 | 67,6 |
| 9 | Menjelaskan Penggunaan Alat dan Bahan | 58,1 | 67,9 | 77,5 |
| 10 | Mengidentifikasi Variabel | 43,5 | 44,8 | 57,8 |
| **Rerata Persentase (%)** | | 50,3 | 53,7 | 69,0 |

Berdasarkan Tabel 3 menunjukkan bahwa dari berbagai kategori sekolah, aspek kognitif proses memprediksi dan mengidentifikasi variabel menunjukkan hasil persentase yang rendah. Hal ini disebabkan rerata siswa dari sekolah yang dijadikan sampel

Copyright © 2020, JPMS, p-ISSN: 1410-1866, e-ISSN: 2549-1458

pengukuran jarang melakukan pembelajaran terkait memprediksi dan mengidentifikasi variabel. Hal ini didukung oleh hasil wawancara dengan guru IPA yang mengajar pada saat kegiatan penelitian. Hasil pengukuran aspek kognitif proses mengkomunikasikan hasil menunjukkan hasil tertinggi. Hal ini disebabkan karena rerata siswa sering melakukan kegiatan terkait dengan mengkomunikasikan hasil kerjanya, seperti persentasi di depan kelas, dan mengubah bentuk penyajian data, serta siswa sering dilatih menyimpulkan hasil kegiatan praktikum dalam lembar kerja siswa. Persentase aspek kognitif proses sekolah berkategori rendah, sedang, dan tinggi secara berturut-turut yaitu 50,3%; 53,7%; dan 69,0%. Sementara itu, untuk lebih jelasnya disajikan pada histogram seperti Gambar 5 berikut.




**Gambar 5.** Perbandingan aspek kognitif proses berdasarkan kategori sekolah

Berdasarkan Gambar 5 dapat diketahui bahwa sekolah yang memiliki kategori tinggi memiliki aspek kognitif proses yang lebih tinggi daripada sekolah kategori sedang. Demikian juga pada sekolah yang memiliki kategori sedang memiliki aspek kognitif proses yang lebih tinggi daripada sekolah dengan kategori sedang. Secara umum kemampuan aspek kognitif pada siswa tergolong baik. Hasil pengukuran aspek kognitif proses dapat diketahui sekolah yang memiliki kategori tinggi cenderung memiliki penguasaan aspek kognitif proses yang lebih tinggi daripada kategori yang sedang dan kategori rendah. Hal ini sesuai dengan hasil wawancara terhadap guru IPA yaitu pada pembelajaran yang dilakukan tiap sekolah berdasarkan kategori sekolah, pada kategori tinggi kegiatan praktikum sering dilakukan dan ketika pembelajaran IPA sering dilakukan di laboratorium.

Sekolah dengan kategori sedang dan rendah kegiatan praktikum dalam satu semester jarang dilakukan. Berdasarkan hasil tersebut menunjukkan sekolah yang sering melakukan pembelajaran IPA dengan kegiatan praktikum memiliki penguasaan aspek kognitif proses yang lebih baik dibandingkan dengan sekolah yang jarang menerapkannya. Hal tersebut sesuai dengan penelitian yang dilakukan Jack (2013) yang mengungkapkan bahwa seks, lokasi sekolah, dan jenis sekolah tidak mempengaruhi perolehan keterampilan proses sains. Sementara itu, kecukupan laboratorium, ukuran kelas, dan pengelolaannya mempengaruhi keterampilan proses sains pada siswa. Hasil pengukuran tersebut sesuai penelitian Rosidin dan Suyatna (2019) yang menemukan bahwa siswa yang mendapatkan skor tes keterampilan proses rendah secara relatif jarang melakukan aktivitas pembelajaran yang berbasis laboratorium.

Hal tersebut membuktikan bahwa aktivitas laboratorium dapat digunakan sebagai media pengembangan keterampilan proses sains. Penelitian lain yang dilakukan oleh Muliyani et al. (2017) juga menemukan bahwa siswa yang telah dilatih keterampilan proses sains memiliki keberhasilan yang lebih baik daripada siswa yang dilatih secara tradisional. Hasil ini menunjukkan bahwa memberikan pelatihan keterampilan proses sains dapat meningkatkan prestasi akademik siswa. Faktor yang turut berpengaruh terhadap kemampuan kognitif proses yaitu secara teoritis sekolah yang termasuk kategori tinggi memiliki kemampuan penguasaan kognitif yang lebih baik daripada dengan sekolah yang memiliki kategori sedang. Selain itu, sekolah dengan kategori sedang secara teoritis memiliki kemampuan penguasaan konsep IPA yang lebih tinggi daripada sekolah dengan kategori rendah.

Berdasarkan beberapa hal itulah yang menjadikan perbedaan yang berbeda pada setiap kategori sekolah. Hasil respon guru terhadap instrumen *integrated assessment* yang dikembangkan memberikan informasi terkait dengan tingkat kelayakan produk instrumen *integrated assessment* yang telah diuji cobakan untuk mengukur aspek kognitif proses pada siswa. Berdasarkan hasil analisis ketiga responden guru di Sleman dihasilkan data tingkat kelayakan produk instrumen *integrated assessment* yang secara umum berkategori baik.

Copyright © 2020, JPMS, p-ISSN: 1410-1866, e-ISSN: 2549-1458

## SIMPULAN

Instrumen *integrated assessment* antara keterampilan proses sains dan penguasaan IPA yang dikembangkan layak untuk digunakan. Penilaian validator diperoleh validitas isi V Aiken 0,98 dan diperoleh 46 butir soal yang valid tanpa revisi dan 4 butir soal valid dengan revisi. Berdasarkan hasil uji coba secara empiris diperoleh 49 butir soal *fit* dan 1 butir soal tidak *fit*. Reliabilitas soal diperoleh nilai 0,73 yang artinya reliabilitas soal tes sudah cukup baik untuk penelitian dasar, sedangkan nilai reliabilitas sampel sangat tinggi karena nilainya 0,92. Hal ini berarti sampel untuk uji coba semakin baik untuk memberikan informasi yang diharapkan. Hasil pengukuran menunjukkan bahwa sekolah yang berkategori tinggi cenderung memiliki penguasaan aspek kognitif proses yang lebih tinggi daripada sekolah dengan kategori sedang dan rendah.

Perbedaan hasil dari pengukuran disebabkan perbedaan sekolah terhadap intensitas pembelajaran di laboratorium dan pengelolaannya sebagai media pengembangan keterampilan proses sains. Faktor yang turut berpengaruh adalah perbedaan kemampuan penguasaan kognitif siswa. Hasil respon keterpakaian produk instrumen *integrated assessment* di lapangan dinyatakan dalam kriteria baik. Dengan demikian, guru IPA dapat menerapkan produk instrumen *integrated assessment* untuk mengukur aspek kognitif proses siswa. Selanjutnya, produk *integrated assessment* dapat digunakan sebagai contoh oleh guru IPA SMP untuk mengembangkan instrumen aspek kognitif proses. Hasil tes ini diperoleh potret kemampuan aspek kognitif proses pada siswa SMP yang dapat digunakan untuk keperluan kebijakan lebih lanjut, seperti pembelajaran IPA berbasis laboratorium untuk meningkatkan *skills*.

## DAFTAR PUSTAKA

Arifin, Z. (2013). *Evaluasi pembelajaran: Prinsip, teknik, dan prosedur*. Remaja Rosdakarya Offset.

Arini, F. Z. R., Susilaningsih, E., & Dewi, N. K. (2017). Pengembangan instrumen penilaian proses untuk mengukur keterampilan sains dan aktivitas siswa SMP. *Journal of Innovative Science Education, 6*(2), 170-178.

Darmayanti, N. W. S., Sadia, W., & Sudiatmika, A. A. I. A. R. (2013). Pengaruh model *collaborative teamwork learning* terhadap keterampilan proses sains dan pemahaman konsep ditinjau dari gaya kognitif. *Jurnal Pendidikan dan Pembelajaran IPA Indonesia, 3*(1), 7-15.

Hasyim, F. (2015). Mengukur kemampuan berpikir analitis siswa SMA. *Jurnal Riset Pendidikan, 1*(02), 121-132.

Jack, G. U. (2013). The influence of identified student and school variables on students' science process skills acquisition. *Journal of Education and Practice, 4*(5), 16-22.

Leighton, J. P., & Gierl, M. J. (2011). *The learning sciences in educational assessment: The role of cognitive models*. Cambridge University Press.

Menteri Pendidikan dan Kebudayaan Republik Indonesia. (2013). *Peraturan Menteri Pendidikan dan Kebudayaan Republik Indonesia nomor 66, tahun 2013, tentang standar penilaian pendidikan.*

Menteri Pendidikan dan Kebudayaan Republik Indonesia. (2013a). *Peraturan Menteri Pendidikan dan Kebudayaan Republik Indonesia Nomor 68, Tahun 2013, tentang kerangka dasar dan struktur kurikulum sekolah menengah pertama/madrasah tsanawiyah.*

Muliyani, R., Kurniawan, Y., & Sandra, D. A. (2017). Peningkatan keterampilan proses sains terpadu siswa melalui implementasi *levels of inquiry* (LoI). *Tadris: Jurnal Keguruan dan Ilmu Tarbiyah, 2*(2), 81-86.

Rosidin, U., & Suyatna, A. (2019). Pengaruh penerapan instrumen *performance assessment* pada pembelajaran IPA berbasis laboratorium real terhadap hasil belajar siswa. *Jurnal Penelitian Pendidikan IPA, 6*(1), 17-26.

Salsabila, K., & Firdaus, A. H. (2018). Pendidikan akhlak menurut Syekh Kholil Bangkalan. *Jurnal Penelitian Pendidikan Islam, 6*(1), 39-56.

Saputro, A. D., Rohaeti, E., & Prodjosantoso, A. K. (2019). Using inquiry-based laboratory instruction to improve critical thinking and scientific process skills among preservice elementary teachers. *Eurasian Journal of Educational Research, 19*(80), 151-170.

Sardinah, S., Tursinawati, T., & Noviyanti, A. (2012). Relevansi sikap ilmiah siswa

Copyright © 2020, JPMS, p-ISSN: 1410-1866, e-ISSN: 2549-1458

dengan konsep hakikat sains dalam pelaksanaan percobaan pada pembelajaran IPA di SDN Kota Banda Aceh. *Jurnal Serambi Ilmu, 13*(2), 70-80.

Shahali, E. H. M., & Halim, L. (2010). Development and validation of a test of integrated science process skills. *Procedia-Social and Behavioral Sciences, 9*(1), 142-146.

Subagia, I. W., & Wiratma, I. G. (2016). Profil penilaian hasil belajar siswa berdasarkan kurikulum 2013. *JPI (Jurnal Pendidikan Indonesia), 5*(1), 39-55.

Subali, B., & Suyata, P. (2012). *Pengembangan item tes konvergen dan divergen dan penyelidikan validitasnya secara empiris.* Diandra.

Sukarno, A. P., & Hamidah, I. (2013). The profile of science process skill (SPS) student at secondary high school (case study in Jambi). *International Journal of Scientific Engineering and Research (IJSER), 1*(1), 79-83.

Sumintono, B., & Widhiarso, W. (2015). *Aplikasi pemodelan rasch pada assessment pendidikan*. Trim komunikata.

Surapranata, S. (2005). *Analisis, validitas, reliabilitas, dan interpretasi hasil tes: Implementasi Kurikulum 2004.* Remaja Rosdakarya.

Suryandari, E. T. (2013). *Performance assessment* sebagai instrumen penilaian untuk meningkatkan keterampilan proses pada praktikum kimia dasar di tadris kimia. *Jurnal Phenomenon, 3*(2), 19-34.

Thiagarajan, S., Semmel, D. S. & Semmel, M. I. (1974). *Instructional development for training teachers of exceptional children.* Leadership Trining Institute/Special Education. University of Minneasota.

Trianto. (2014). *Model pembelajaran terpadu: Edisi ke-5*. Bumi Aksara.

Usman, M.U. (2006). *Menjadi guru profesional*. Remaja Rosdakarya.

Winata, A., & RW, I. S. (2018). Kemampuan awal literasi sains peserta didik kelas V SDN Sidorejo I Tuban pada materi daur air. *JTIEE (Journal of Teaching in Elementary Education), 2*(1), 58-64.

.

Copyright © 2020, JPMS, p-ISSN: 1410-1866, e-ISSN: 2549-1458