

KARAKTERISASI TES PENALARAN ILMIAH (TPI) PADA MATERI HUKUM GRAVITASI MENGGUNAKAN PENDEKATAN *ITEM RESPONSE THEORY*

Indri Liani Sartika ^{*}), Ridwan Efendi, Taufik Ramlan Ramalis

Departemen Pendidikan Fisika, FPMIPA, Universitas Pendidikan Indonesia

* Email : slianindri04@gmail.com

ABSTRAK

Penalaran ilmiah merupakan keterampilan penting yang seharusnya dimiliki siswa pada abad ke-21. Namun, kemampuan ini di Indonesia masih tergolong rendah. Hal ini terlihat dari hasil tes PISA dimana peringkat Indonesia berada di posisi 64 dari total 72 negara. Salah satu bentuk penanggulangan permasalahan tersebut, peneliti melakukan karakteristik Tes Penalaran Ilmiah (TPI) pada materi hukum gravitasi dengan menggunakan *item response theory*. Tes dilakukan pada siswa SMA kelas X dan XI MIA di lima sekolah negeri di kota Bandung. Metode yang digunakan pada penelitian ini adalah metode kuantitatif dengan desain konstruksi dan validasi. Data dianalisis menggunakan *partial credit model*, 15 dari soal 14 yang dikonstruksi soal memiliki indeks kesukaran antara $-2 \leq b \leq 2$ artinya tergolong dalam kategori baik. Soal reliabel untuk kemampuan siswa rendah, sedang, dan tinggi karena bernilai $-0,1-2$.

Kata Kunci: Penalaran Ilmiah; *Item Response Theory*; *Partial Credit Model*

ABSTRACT

Scientific reasoning is an important skill that students should have in the 21st century. However, this ability in Indonesia is still relatively low. This is evident from the results of the PISA test where Indonesia ranked in position 64 of a total of 72 countries. One form of overcoming these problems, researchers do characteristic of Scientific Reasoning Test (TPI) on gravity law material by using item response theory. The test was conducted on high school students of class X and XI MIA in five public schools in Bandung. The method used in this research is quantitative method with construction design and validation. Data were analyzed using a partial credit model, 15 of which 14 constructed about having an index of difficulty between $-2 \leq b \leq 2$ means belonging to either category. Reliable problem for low, medium, and high student ability because it is worth $-0,1-2$.

Keywords: Scientific Reasoning Test; Item Response Theory; Partial Credit Models

PENDAHULUAN

Penalaran Ilmiah adalah proses dimana prinsip-prinsip logika diterapkan pada proses ilmiah –pencarian mengenai suatu penjelasan, formulasi dari suatu hipotesis, membuat prediksi, solusi dari suatu masalah, penyusunan eksperimen, kontrol variabel, menganalisis data, dan pengembangan hukum empiris-semuanya dalam cara yang logis [1]). Penalaran ilmiah merupakan salah satu aspek kognitif yang seharusnya dimiliki siswa di Indonesia. Hal ini dikarenakan, ditemukan adanya korelasi antara kemampuan penalaran ilmiah siswa dengan kemajuan belajar pada konten sains [2]. Selain itu, beberapa kemampuan penalaran ilmiah dapat membantu dalam pengambilan

keputusan ataupun pemecahan masalah dalam kehidupan sehari-hari. Pada kenyataannya, penalaran ilmiah sangat erat kaitannya dengan kehidupan sehari-hari tiap individu. Seperti penalaran proporsional yang digunakan untuk menentukan harga termurah dari suatu barang, penalaran induktif yang dapat digunakan untuk mengambil kesimpulan ketika informasi yang tersedia sangat terbatas, dan penalaran probabilitas maupun klausul yang dapat digunakan dalam memprediksi cuaca.

Pada kenyataannya kemampuan penalaran ilmiah siswa di Indonesia masih tergolong rendah. Berdasarkan hasil tes PISA pada tahun 2015, Indonesia berada pada peringkat ke-64 dari total 72 negara [3]. Sedangkan untuk skor rata-rata pada bidang sains Indonesia

memperoleh skor 403 poin. Skor tersebut berada dibawah skor rata-rata yang ditetapkan OECD yaitu 493 poin [3]. Tes ini, merupakan suatu bentuk evaluasi kemampuan dan pengetahuan yang dirancang untuk siswa usia 15 tahun pada bidang membaca, matematika, dan sains (Shiel, 2007). Ada tujuh aspek yang digunakan dalam penilaian, salah satunya adalah penalaran ilmiah. Untuk prestasi fisika yang diukur pada aspek penalaran Indonesia berada pada ranking 40 dari 42 negara [4]. Berdasarkan hasil TIMSS dapat disimpulkan bahwa: (1) rata-rata capaian fisika siswa Indonesia ditinjau dari aspek kognitif (*knowing*, *applying*, dan *reasoning*) masih rendah; (2) kecenderungan capaian fisika siswa di Indonesia selalu menurun pada tiap aspek kognitif sehingga kemampuan fisika siswa Indonesia harus ditingkatkan pada semua aspek penalaran dengan cara membekali siswa kemampuan berpikir tingkat tinggi [5].

Salah satu faktor penyebab rendahnya hasil PISA adalah kurang terlatihnya siswa dalam mengerjakan soal-soal berkarakteristik PISA yang terdiri dari soal berlevel tinggi. Hal itu terjadi karena sistem evaluasi di Indonesia yang masih menggunakan soal level rendah. Untuk mengunggulangnya, penelitian ini bermaksud untuk mengembangkan Tes Penalaran Ilmiah (TPI). Sebelumnya, telah ada beberapa ahli yang mengembangkan tes mengenai penalaran ilmiah seperti Lawson yang pada tahun 1978 mengembangkan dan memvalidasi *Lawson Classroom Test of Formal Reasoning* yang didesain untuk mengukur penalaran konkret dan formal pada siswa sekolah menengah dan mahasiswa. Pada tahun 2013, Han dan tim penelitiannya mengembangkan *Inventory for Scientific Thinking and Reasoning (iSTAR) Assessment*. Mereka mengidentifikasi delapan dimensi penalaran ilmiah, yang merupakan pengembangan enam dimensi penalaran ilmiah yang dikembangkan oleh Lawson. Pada penelitian ini, penulis berfokus untuk mengkonstruksi tes penalaran ilmiah merujuk pada i-STAR yang dikembangkan oleh Jing Han.

Penalaran ilmiah merupakan cara berpikir kritis yang melibatkan kemampuan berpikir analisis induktif dan deduktif [6]. Fisika merupakan salah satu cabang sains, yang meliputi rasa ingin tahu tentang benda dan fenomena alam yang menimbulkan masalah baru dan dapat diselesaikan melalui metode ilmiah yang meliputi penyusunan hipotesis, perancangan eksperimen atau percobaan,

evaluasi, pengukuran, dan penarikan kesimpulan [6]. Hakikat fisika tersebut sangat sesuai dengan proses penalaran ilmiah. Berdasarkan pemaparan tersebut, Fisika dapat digunakan sebagai sarana melatih penalaran ilmiah. Akan tetapi, pada kenyataannya, selama pembelajaran peserta didik cenderung untuk menghafal persamaan fisika tanpa mampu memahami konsepnya. Selain itu, soal-soal yang seringkali peserta didik kerjakan hanya sebatas “penerapan” rumus fisika yang diberikan selama pembelajaran tanpa siswa mengetahui maksud fisis dari rumus tersebut. Sehingga, kemampuan siswa dalam hal mengingat persamaan fisika dan kemampuan matematis bertambah namun pengetahuan fisiknya cenderung konstan.

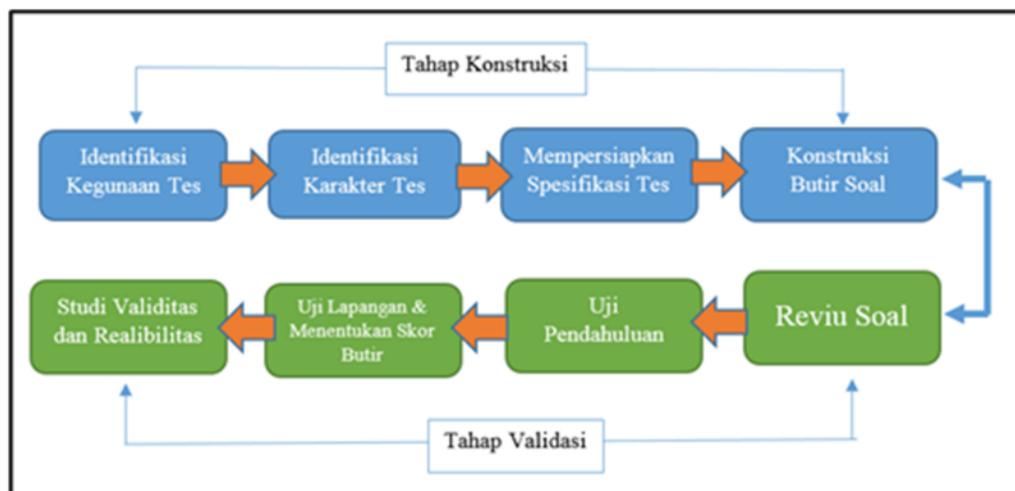
Terdapat berbagai materi fisika mengenai konsep Newtonian, salah satunya mengenai hukum newton tentang gravitasi. Pemilihan materi hukum newton tentang gravitasi dikarenakan pada materi tersebut dibutuhkan pengetahuan awal yang kuat mengenai ketiga hukum newton tentang gerak dan perlunya pemahaman yang tinggi sehingga hukum newton tentang gravitasi dapat dijadikan sarana untuk penggalian penalaran ilmiah pada siswa. Selain itu, penelitian mengenai materi ini belum banyak diteliti. Padahal, hukum newton tentang gravitasi adalah materi yang cukup abstrak dan dikatakan siswa sebagai materi fisika yang sulit.

Penilaian pendidikan adalah proses pengumpulan dan pengolahan informasi untuk menentukan pencapaian hasil belajar peserta didik [7]. Bagi pendidik, penilaian dapat dijadikan acuan dalam menetapkan ketuntasan penguasaan materi, menetapkan seberapa perlunya diadakan program perbaikan atau pengayaan pada suatu materi, dan juga memperbaiki proses pembelajaran. Untuk mendapatkan tujuan-tujuan dari penilaian yang diharapkan, dibutuhkan tes yang baik, yaitu mampu mengukur aspek-aspek dari tujuan yang diharapkan pendidik. Dalam pengelolaan instrumen, terdapat dua jenis analisis secara empirik yaitu *Classical Test Theory (CTT)* dan *Item Response Theory (IRT)*. Namun, pada penelitian ini digunakan pendekatan model *Item Response Theory*, karena pada *Classical Test Theory* parameter butir dan parameter peserta saling tergantung sehingga tidak dapat digeneralisasikan untuk kelompok peserta didik yang lain. Selain itu, pada CTT, sebuah soal yang diberikan pada peserta didik dengan kemampuan tinggi maka tingkat kesukarannya akan menjadi kecil, begitu juga sebaliknya, jika

soal yang diberikan pada siswa yang berkemampuan rendah maka tingkat kesukarannya akan menjadi tinggi. Penggunaan IRT mampu menanggulangi masalah-masalah tersebut.

Penelitian menggunakan metode kuantitatif dengan desain konstruksi dan validasi yang dikembangkan oleh Crocker dan Algina. Tahapan dan langkah-langkah pada desain konstruksi dan validasi ditunjukkan pada gambar berikut:

METODE PENELITIAN



Gambar 1. Langkah-Langkah pada Desain Konstruksi dan Validasi

Pada tahap konstruksi dilakukan kegiatan identifikasi kegunaan tes, identifikasi karakteristik tes, mempersiapkan spesifikasi tes, dan konstruksi butir. Keseluruhan tahap ini dimaksudkan untuk mendapatkan seperangkat TPI yang layak di uji coba secara luas di lapangan.

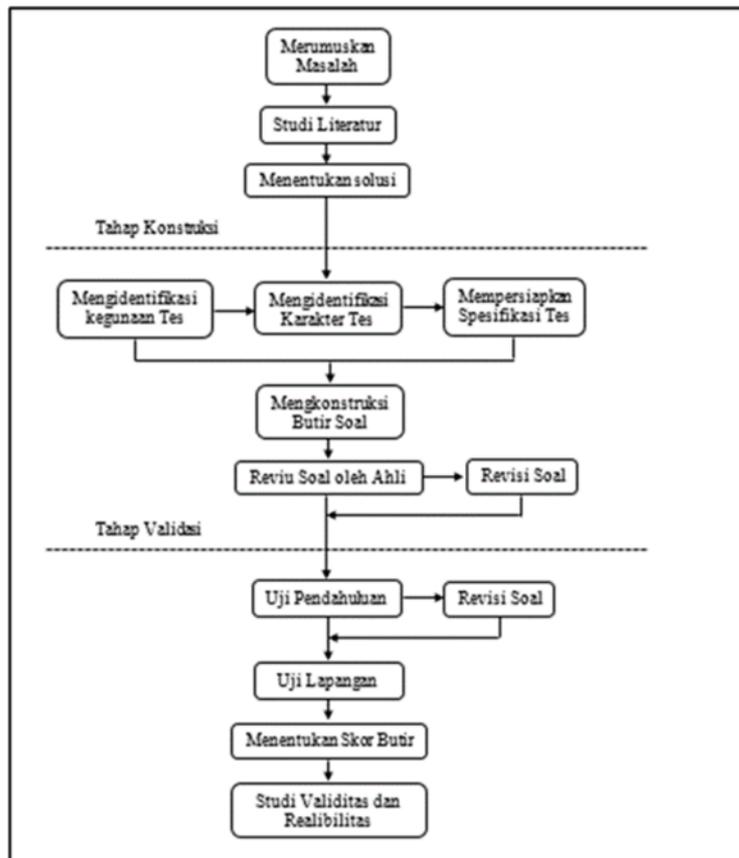
Pada tahap validasi dilakukan kegiatan reviu soal (dilakukan tiga orang ahli di bidang penalaran ilmiah, bumi dan antariksa, dan evaluasi pendidikan), uji pendahuluan, uji lapangan dan menentukan skor butir, dan melakukan validasi dan realibilitas tes. Tujuan dilakukan tahap ini adalah untuk mendapatkan dan mengolah data sehingga bisa diperoleh informasi yang sesuai dengan tujuan penelitian.

Subjek penelitian adalah siswa kelas X dan XI MIA yang telah mempelajari hukum newton tentang gravitasi. Uji lapangan (*field test*) dilakukan pada lima

sekolah berbeda di kota Bandung dengan karakteristik yang beraneka ragam. Tes

dilakukan pada dua tingkatan kelas yang berbeda, yaitu pada siswa kelas X MIA dan XI MIA. Hal ini dilakukan karena pada kurikulum nasional materi mengenai hukum newton tentang gravitasi diberikan pada kelas X MIA di semester genap. Sedangkan pada kurikulum 2013, materi disampaikan pada siswa kelas XI MIA di semester ganjil.

Penelitian ini menggunakan instrumen soal pilihan ganda disertai alasan yang juga berbentuk pilihan ganda. Sumber tipe instrumen berasal dari peneliti lain yang telah mengembangkan tes jenis ini. Kemudian penulis membuat instrumen dengan tipe serupa namun dengan materi berbeda yaitu mengenai hukum gravitasi. Soal di konstruksi berdasarkan delapan aspek penalaran ilmiah yang dikembangkan oleh Jing Han.



Gambar 2. Langkah-Langkah Karakterisasi TPI mengenai Hukum Gravitasi

Analisis statistik pada penelitian ini menggunakan bantuan aplikasi *IRTPRO4 student*. Adapun model IRT yang digunakan adalah model *Partial Credit Model (PCM)*. Model ini digunakan karena sesuai dengan tipe tes yang dikembangkan yaitu pada penskorannya terdapat empat kemungkinan skor yaitu 0, 25, 75, dan 100. Item dikatakan baik jika indeks kesukaran lebih dari -2 atau kurang dari 2. Berdasarkan fungsi informasi dan SEM, maka dapat diketahui bahwa tes cocok untuk siswa dengan kemampuan (θ) tertentu dari sangat rendah hingga sangat tinggi.

HASIL DAN PEMBAHASAN

1. Hasil Konstruksi Tes

TPI terdiri 15 soal mengenai hukum gravitasi yang dikonstruksi berdasarkan 8 aspek penalaran ilmiah yang dikembangkan oleh Jing Han yaitu variabel kontrol, penalaran proporsional, penalaran probabilitas, penalaran deduktif, penalaran induktif, penalaran klausal, penalaran korelasional, dan penalaran hipotetif-deduktif.

2. Hasil Validasi

Proses reviu soal oleh tiga orang ahli di bidang penalaran ilmiah, bumi dan antariksa, dan evaluasi pendidikan. Berdasarkan kegiatan ini, ketiga validator mengatakan soal telah layak digunakan. Setelah itu, dilakukan uji pendahuluan pada 6 orang siswa yang mewakili subjek pada uji lapangan. Peserta percobaan awal dipersilakan mengisi soal yang telah disusun, kemudian pengembang tes mengamati perilaku yang terlihat pada peserta tes seperti jeda berfikir panjang atau perubahan jawaban pada butir tertentu. Melalui tahap ini, pengembang tes dapat menentukan berapa lama waktu yang diperlukan peserta untuk mengerjakan tes pada uji lapangan melalui pola waktu penyelesaian tes peserta pada tahap ini. Setelah mengerjakan soal, pengembang tes melakukan wawancara dengan peserta tes untuk mengetahui komentar peserta tes pada perangkat soal yang telah di konstruksi. Pada langkah ini, peneliti berhasil memperbaiki kesalahan kata pada tiga soal yaitu nomor 9, 12, dan 15. Selain itu, peneliti juga menambahkan keterangan kata terhadap istilah-istilah yang kurang dipahami oleh siswa yaitu pada nomor 13 dan 14. Pada langkah ini pula, peneliti

memutuskan waktu pengerjaan soal untuk TPI pada materi hukum gravitasi pada uji lapangan selama 90 menit atau 2xjam pelajaran.

Uji lapangan dilakukan selama lima hari di lima sekolah yang berbeda. SMAN A pada Senin, 20 Maret 2017. SMAN B pada Rabu, 15 Maret 2017. SMAN C pada Rabu, 5 April 2017. SMAN D pada Selasa, 28 Maret 2017. SMAN E pada Rabu, 29 Maret 2017. Pada uji lapangan, peneliti mendapatkan sebanyak 158 data dari kelima sekolah tersebut. Adapun, data tersebut didapatkan dari kelas yang berbeda. Pada SMAN A, SMAN D, dan SMAN E data diambil pada siswa/i kelas X MIA. Sedangkan, pada SMAN B dan SMAN C data diambil pada siswa/i kelas XI MIA. Hal ini terjadi karena penerapan kurikulum yang berbeda pada kelas yang berbeda pada SMA di kota Bandung. Pada Kurikulum Nasional (Kurnas) materi Hukum Newton tentang Gravitasi diajarkan pada kelas X MIA di semester genap. Sedangkan, pada Kurikulum 2013, materi ini diajarkan pada kelas XI MIA di semester ganjil.

3. Kurva Karakteristik Butir

Studi validitas dan reabilitas adalah tahap terakhir dari keseluruhan tahap pada penelitian ini. Skor yang telah di *input* pada aplikasi *IRTPRO4 Student* akan menghasilkan kurva karakteristik butir untuk tiap butirnya.

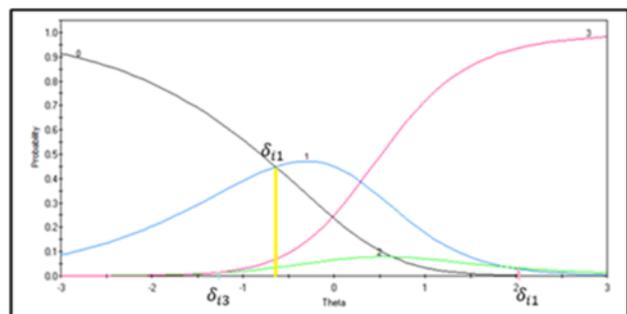
Karakteristik butir ditunjukkan dengan kurva karakteristik butir dan indeks kesukaran. Berdasarkan analisis diperoleh 15 kurva karakteristik butir. Pada gambar 3 disajikan contoh kurva karakteristik butir pada butir 1, yang dapat dijelaskan bahwa:

- Skor 0 (kategori 0) sebagian besar diperoleh responden dengan kemampuan sangat rendah yaitu pada nilai θ sebesar -3;
- Skor 25 (kategori 1) sebagian besar diperoleh responden dengan kemampuan rendah yaitu pada nilai θ sebesar -0,3;
- Skor 75 (kategori 2) sebagian besar diperoleh responden dengan kemampuan rendah yaitu pada nilai θ sebesar 0,4;
- Skor 100 (kategori 3) sebagian besar diperoleh responden dengan kemampuan tinggi yaitu pada nilai θ sebesar 3.

Berdasarkan informasi kemampuan dari tiap kategori dapat diurutkan tingkat kesulitan kategori dari kecil ke besar secara berurutan adalah kategori 0, 1, 2, dan 3.

δ_{ij} menunjukkan titik perpotongan dua garis probabilitas kategori berurutan dalam satu butir. Melalui grafik gambar 3 diperoleh informasi sebagai berikut:

- Perpotongan garis probabilitas kategori 0 dan 1 (δ_{i1}) menunjuk nilai θ sebesar -0,6 artinya untuk memperoleh kategori 1 dari kategori 0 dibutuhkan kemampuan -0,6. Sehingga dapat dikatakan tingkat kesulitan untuk bertransisi dari kategori 0 ke kategori 1 adalah -0,6 yang termasuk ke dalam kategori sedang.
- Perpotongan garis probabilitas kategori 1 dan 2 (δ_{i2}) menunjuk nilai θ sebesar 2 artinya untuk memperoleh kategori 2 dari kategori 1 dibutuhkan kemampuan 2. Sehingga dapat dikatakan tingkat kesulitan untuk bertransisi dari kategori 1 ke kategori 2 adalah 2 yang termasuk ke dalam kategori sukar.
- Perpotongan garis probabilitas kategori 2 dan 3 (δ_{i3}) menunjuk nilai θ sebesar -1,2 artinya responden yang telah memperoleh skor 75 (kategori 2) membutuhkan kemampuan minimal -1,2 untuk dapat memperoleh skor 100 (kategori 3). Sehingga dapat dikatakan tingkat kesulitan untuk bertransisi dari kategori 2 ke kategori 3 adalah -1,2 yang termasuk ke dalam kategori sedang.



Gambar 3. Kurva Karakteristik Butir 1

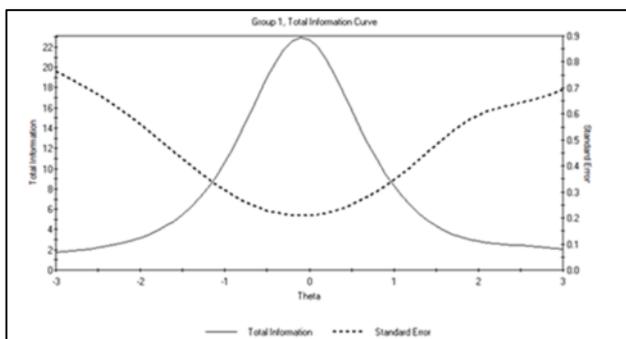
Berdasarkan informasi mengenai kemampuan pada tiap perpotongan kategori yang berurutan maka dapat diurutkan tingkat kesulitan δ_{ij} dari kecil ke besar adalah δ_{i3} , δ_{i2} , dan δ_{i1} .

No. Soal	Kesulitan Butir	Kriteria Butir	Kesulitan Kategori				Kesulitan δ_{ij}		
			1	2	3	4	δ_{i1}	δ_{i2}	δ_{i3}
1	-0.05	Baik	-3	-0,3	0,4	3	-0,6	2	-1,2
2	-0.19	Baik	-3	-0,45	0,2	3	2,2	0,2	-1,7
3	0.01	Baik	-3	-0,25	0,35	3	0,6	2,6	-
4	2.66	Kurang Baik	-3	-	2,65	3	-	-	-
5	-0.54	Baik	-3	-0,8	0,5	3	0,4	1,3	-2,80
6	-0.46	Baik	-3	-0,7	0,1	3	-1,67	2	-2
7	0.36	Baik	-3	0,1	0,65	3	1,1	1,1	-1,2
8	0.14	Baik	-3	-0,1	0,4	3	1,6	-0,01	-1,4
9	-0.18	Baik	-3	-0,6	-0,1	3	1,8	-0,01	-1,4
10	-0.16	Baik	-3	-0,5	0,2	3	0,6	0,3	-1,3
11	-0.83	Baik	-3	-1,3	-0,5	3	-0,5	-1,1	-1
12	-1.04	Baik	-3	-1,2	0,4	3	-2,7	1,75	-
13	-0.27	Baik	-3	-0,6	0,3	3	0,7	-0,2	-1,7
14	0.94	Baik	-3	1	1,8	3	2,6	0,4	1,4
15	-0.28	Baik	-3	-1,6	-0,2	3	1,1	0,7	-2,3

Gambar 4. Distribusi Tingkat Kesulitan TPI mengenai Hukum Gravitasi

4. Fungsi Informasi dan SEM

Hasil analisis diperoleh fungsi informasi dan *standar error measurement*.



Gambar 5. Kurva Informasi Total TPI pada Materi Hukum Gravitasi

Berdasarkan gambar 5, kurva TPI memiliki nilai maksimum 60 pada skala kemampuan 0,2. Perpotongan antara kurva fungsi informasi dengan kesalahan pengukuran berada pada rentang -0,1 hingga 1,2. Hal ini menunjukkan bahwa TPI pada materi hukum gravitasi dapat digunakan untuk mengukur kemampuan pada rentang -0,1 sampai 1,2 yang reliabel untuk mengukur kemampuan rendah, sedang, hingga tinggi.

SIMPULAN DAN SARAN

Dalam pelaksanaan penelitian ini dilakukan tahap konstruksi dan validasi guna mendapatkan karakteristik mengenai Tes

Penalaran Ilmiah (TPI) mengenai hukum gravitasi.

1. Tahap konstruksi dilakukan untuk menghasilkan TPI yang sesuai dengan standar TPI yang sebelumnya telah dikembangkan oleh Jing Han. Untuk mencapai tujuan tersebut dilakukan empat langkah kegiatan yaitu identifikasi kegunaan tes, identifikasi karakter tes, mempersiapkan spesifikasi tes, dan konstruksi butir soal. Melalui tahapan ini, dihasilkan 15 butir soal *multiple choice* yang mewakili delapan dimensi penalaran ilmiah.
2. Tahap validasi dilakukan untuk mengkarakteristik TPI mengenai hukum gravitasi. Dalam upaya mencapai tujuan tersebut dilakukan langkah kegiatan sebagai berikut: reviu soal, percobaan awal, uji lapangan dan menentukan nilai butir, dan melakukan studi validitas dan realibilitas.
3. Keempat belas soal dalam keadaan baik karena tingkat kesulitan butirnya berada di rentang -2 hingga 2. Namun, ada 1 soal pada nomor 4 mengenai penalaran proporsional tingkat kesulitannya bernilai 2,26 artinya soal termasuk dalam kategori sukar.
4. Selain karakterisasi pada tiap soalnya, melalui grafik fungsi informasi didapatkan bahwa TPI mengenai hukum gravitasi yang dikonstruksi dapat digunakan untuk mengukur kemampuan pada rentang -0,1 hingga 1,2 artinya pada kemampuan rendah hingga tinggi.

Saran dalam penelitian ini adalah sebagai berikut:

1. Instrumen tes ini direkomendasikan untuk digunakan di sekolah karena instrumen tes ini dapat mengukur kemampuan penalaran ilmiah siswa pada rentang -0,1 hingga 4.
2. TPI dapat dikembangkan pada materi fisika lain selain materi mengenai hukum gravitasi.
3. TPI dapat dilakukan penelitian lebih lanjut dengan menggunakan analisis data politomus menurut *generalized partial credit model* (GPCM 3PL).

REFERENSI

- [1] Wenning, C. J. & Vierya, R. (2015). *Teaching High School Physics*, 1. Publisher: Authors.
- [2] Lawson, A. E. (2005). *What Is the Role of Induction and Deduction in Reasoning and Scientific Inquiry Journal of Research in Science Teaching*, 42 (6), 716-740.
- [3] PISA 2015 Result in Focus. (2015). *Programme for International Student Assessment*. Diambil Tanggal 12 Desember 2016, dari <http://www.oecd.org/pisa>.
- [4] TIMSS & PIRLS International Study Center. (2012). *TIMSS 2011 international result in science*. Boston: The TIMSS & PIRLS International Study Center, Boston College. Diambil tanggal 29 Mei 2017, dari <http://timss.bc.edu/timss2011/release.htm> /
- [5] Efendi, R. (2010). Kemampuan Fisika Siswa Indonesia dalam TIMSS. *Prosiding Seminar Nasional Fisika 2010*. ISBN: 978-979-98010-6-7.
- [6] Erlina, N, dkk. (2016). Penalaran Ilmiah dalam Pembelajaran Fisika. *Research Gate*, 473-479. ISBN: 978-602-72071-1-0.
- [7] Depdiknas. (2007). *Peraturan Menteri Pendidikan Nasional No 20 Tahun 2007 tentang Standar Penilaian*.