

BAB IV

HASIL PENELITIAN DAN PEMBAHASAN

A. Hasil Pengembangan Produk Awal

Produk pengembangan yang dibuat dalam penelitian ini adalah soal tes HOTS mata pelajaran IPS kelas VIII semester 2. Produk soal tes HOTS yang dibuat berbentuk pilihan ganda dengan alternatif jawaban empat (a,b,c,d) dan soal uraian. Jumlah soal keseluruhan 30 butir soal yang terdiri dari 26 pilihan ganda dan 4 uraian. Langkah-langkah pengembangan dalam penelitian ini mengacu pada prosedur dan pengembangan yang dikemukakan oleh Borg & Gall. Namun tetap disesuaikan dengan kepentingan dan tujuan penelitian. Adapun tahapan penelitian dan pengembangan instrumen HOTS dalam penelitian ini adalah: (1) survei lapangan, (2) perencanaan, (3) pembuatan produk instrumen penilaian HOTS, (4) melakukan validasi dosen ahli dan revisi dosen ahli, (5) uji coba lapangan awal dan revisi awal, (6) uji coba akhir, (7) analisis hasil uji coba dan penyempurnaan produk utama.

Untuk dapat melihat tingkat kualitas produk pengembangan instrumen penilaian HOTS peneliti disini menggunakan dua kriteria penilaian yakni valid dan reliabel. Instrumen penilaian HOTS yang sudah dibuat harus melewati dua tahap penilaian tersebut. Tahap penilaian pertama yaitu untuk melihat kevalidan instrumen penilaian HOTS. Pada tahap kevalidan disini menggunakan penilaian dari dosen ahli. Tahap penilaian kedua yaitu untuk melihat reliabel instrumen penilaian HOTS. Pada tahap reliabel disini dilakukan uji coba lapangan di SMP Negeri 5 Yogyakarta kelas VIII yang berjumlah 288 siswa. Pada tahap ini penilaian difokuskan pada reliabilitas,

tingkat kesukaran butir soal, daya beda butir soal, dan keefektifan distraktor/ pengecoh. Proses dalam pengembangan instrumen penilaian HOTS disini meliputi berdiskusi dengan guru untuk menentukan SK & KD yang akan digunakan dalam menyusun butir-butir soal tes HOTS. Setelah SK & KD sudah disepakati oleh peneliti dan guru, maka selanjutnya peneliti membuat kisi-kisi soal dan menyusun instrumen penilaian HOTS.

Instrumen penilaian HOTS yang sudah selesai dirancang untuk selanjutnya diserahkan kepada dosen validator untuk dilakukan penilaian. Hasil dari penilaian dosen validator tersebut sebagai bahan untuk dilakukan revisi yang selanjutnya di peroleh produk awal yang sudah siap untuk digunakan sebagai bahan uji coba lapangan awal. Uji coba lapangan awal dilakukan pada skala yang lebih kecil sehingga peneliti melakukan uji coba pada satu kelas yaitu kelas VIII H. Setelah selesai melakukan tahap uji coba lapangan awal maka diperoleh hasil analisis untuk dilakukan revisi sehingga akan menjadi produk utama instrumen penilaian HOTS yang siap digunakan untuk uji coba lapangan. Hasil dari uji coba lapangan tersebut yang nantinya dapat dilihat tingkat kemampuan HOTS siswa SMP Negeri 5 Yogyakarta. Selain itu juga diperoleh data estimasi koefisien reliabilitas, tingkat kesukaran butir soal, daya beda butir soal, dan keefektifan distraktor/ pengecoh instrumen penilaian HOTS.

B. Hasil Uji Coba Produk

1. Data Hasil Validasi Ahli

Validitas isi sangat ditekankan guna memperoleh kualitas instrumen penelitian yang baik sehingga validasi dalam penelitian dilakukan dengan telaah kualitatif dan kuantitatif. Untuk telaah kuantitatif menggunakan analisis formula Aiken's V untuk menghitung *content validity coefficient*, namun sebelum pada tahap analisis formula Aikens's V terlebih dahulu validasi dilakukan oleh dosen ahli, dalam penelitian ini dosen validator yaitu Prof. Sudji Munadi, M.Pd, Dr. Supardi, S.Pd., M.Pd, dan Saefur Rochmat, MIR., Ph.D. Kegiatan validasi ini untuk mengetahui layak atau tidaknya instrumen penilaian HOTS yang sudah dibuat. Kegiatan validasi bukan hanya instrumen penilaian HOTS tetapi juga kuesioner HOTS untuk siswa.

Hasil dari validasi akan diperoleh masukan, saran, dan penilaian instrumen penilaian HOTS dan kuesioner. Proses pada tahapan validasi ini dilakukan dengan menyerahkan produk pengembangan yang berupa kisi-kisi soal instrumen penilaian HOTS, dan kuesioner, serta lembar penilaian yang harus diisi oleh dosen validator. Setelah proses validasi selesai dilakukan oleh kedua dosen, maka selanjutnya adalah menganalisis hasil validasi pada setiap butir soal dan kuesioner. Berikut ini adalah data hasil analisis validasi ahli menggunakan formula Aiken's V dapat dilihat pada tabel 14.

Tabel. 14 Hasil Analisis Validasi Ahli Soal HOTS Pilihan Ganda

Nomor Butir Soal	Koefisien Aiken's V	Kriteria
1	0,83	Layak digunakan
2	0,91	Layak digunakan
3	0,91	Layak digunakan
4	0,83	Layak digunakan
5	0,83	Layak digunakan
6	1,00	Layak digunakan
7	0,83	Layak digunakan
8	1,00	Layak digunakan
9	0,83	Layak digunakan
10	0,91	Layak digunakan
11	0,83	Layak digunakan
12	0,75	Layak digunakan
13	0,83	Layak digunakan
14	0,83	Layak digunakan
15	0,91	Layak digunakan
16	0,75	Layak digunakan
17	0,83	Layak digunakan
18	0,91	Layak digunakan
19	1,00	Layak digunakan
20	1,00	Layak digunakan
21	0,83	Layak digunakan
22	0,91	Layak digunakan
23	0,91	Layak digunakan
24	0,75	Layak digunakan
25	1,00	Layak digunakan
26	1,00	Layak digunakan

Tabel. 15 Hasil Analisis Validasi Ahli Soal HOTS Uraian

Nomor Butir Soal	Koefisien Aiken's V	Kriteria
1	0,83	Layak digunakan
2	0,91	Layak digunakan
3	0,75	Layak digunakan
4	0,91	Layak digunakan

Berdasarkan dari hasil analisis menggunakan formula Aiken's V menunjukkan semua soal pilihan ganda yang berjumlah 26 dan uraian 4 semuanya dikatakan layak untuk digunakan. Namun perlu diperhatikan

bahwa, terdapat beberapa soal yang harus di perbaiki guna untuk memperoleh kualitas soal yang lebih baik. Perbaikan instrumen penilaian dilakukan sesuai dari masukan dan saran dari ketiga dosen validator. Untuk telaah butir soal secara kualitatif digunakan berdasarkan penilaian *judgment* dari ahli dengan melihat aspek materi, konstruksi, dan bahasa. Telaah butir soal secara kualitatif didasarkan pada ketentuan sebagai berikut:

- a. Baik, apabila keseluruhan kriteria telaah butir pada aspek materi, konstruksi dan bahasa semua sesuai dengan kaidah dan didukung oleh semua penelaah (81-100%).
- b. Kurang baik, apabila butir tersebut tidak memenuhi sebanyak-banyaknya satu aspek materi, tiga aspek konstruksi serta satu aspek bahasa (70-80%).
- c. Tidak baik, apabila butir tidak memenuhi semua kriteria yang telah ditetapkan (< 69%).

Berdasarkan hasil telaah secara kualitatif pada 30 butir soal, maka secara keseluruhan disajikan bentuk persentase pada tabel 16.

Tabel. 16 Hasil Telaah Kulitatif

Aspek Penelaahan	Kriteria Penilaian	Validator			Ket
		1	2	3	
Materi	Butir soal sesuai dengan indikator	90%	90%	90%	Baik
	Pilihan jawaban logis	80%	85%	80%	Baik
	Hanya ada kunci jawaban yang paling tepat	85%	90%	82%	Baik
	Materi yang diukur sesuai dengan kompetensi yang diukur	90%	84%	90%	Baik
Konstruksi	Soal dirumsukan secara jelas	85%	75%	75%	Baik
	Rumusan pokok soal dan pilihan jawaban merupakan pernyataan yang diperlukan	75%	90%	70%	Baik
	Soal tidak memberi petunjuk kunci jawaban	80%	82%	80%	Baik
	Soal bebas dari pernyataan yang bersifat negatif	85%	85%	83%	Baik
	Pilihan jawaban logis ditinjau dari segi materi	90%	83%	75%	Baik
	Gambar, grafik, tabel diagram berfungsi jelas	83%	95%	78%	Baik
	Panjang pilihan jawaban tidak menggunakan pernyataan semua jawaban di atas salah/benar dan sejenisnya	83%	95%	80%	Baik
	Pilihan jawaban yang berbentuk angka atau waktu disusun berdasarkan urutan besar kecilnya angka atau kronologis	82%	90%	85%	Baik
	Butir soal tidak bergantung pada jawaban sebelumnya	75%	90%	85%	Baik
	Bahasa	Kesesuaian dengan kaidah bahasa Indonesia	78%	85%	80%
Soal menggunakan bahasa yang komunikatif		80%	80%	78%	Baik
Soal tidak menggunakan bahasa yang tabu		90%	90%	85%	Baik

Berdasarkan dari tabel 16 terlihat bahwa rata-rata instrumen penilaian HOTS menunjukkan angka yang cukup baik, hanya saja ada beberapa

butir-butir soal yang memerlukan perbaikan. Aspek materi rata-rata berada pada angka 90% yang berarti menunjukkan kategori baik karena kisi-kisi dan indikator soal sesuai dengan apa yang hendak diukur. Pada aspek konstruksi rata-rata berada pada angka 85% yang berarti juga menunjukkan kategori baik. Namun ada beberapa yang perlu diperhatikan yaitu mengenai butir-butir soal yang menggunakan gambar. Gambar yang digunakan usahakan dapat berfungsi untuk menjawab butir soal. Selain itu ada beberapa soal yang rumusan soalnya masih kurang jelas antara stem soal dan pilihan jawaban.

Hal ini dapat menimbulkan persepsi ganda pada siswa, maka untuk itu perlu diperbaiki. Aspek bahasa secara keseluruhan rata-rata berada pada angka 80% yang berarti menunjukkan kategori kurang baik. Ada yang perlu diperhatikan dalam penggunaan huruf kapital dan SPOK. Dapat disimpulkan bahwa dari masing-masing aspek sudah menunjukkan kategori baik, hanya sedikit perlu revisi dibagian gambar, rumusan soal, dan stem soal dan jawaban yang masih rancu. Hasil telaah validasi instrumen penilaian HOTS ini sebagai dasar bagi peneliti untuk memperbaiki instrumennya sebelum nantinya benar-benar dapat digunakan dilapangan. Proses dalam melakukan revisipun sesuai dari masukan dan saran dosen validator. Setelah selesai melakukan revisi instrumen penilaian HOTS dikembalikan lagi pada dosen validator untuk dapat dilihat kembali apakah instrumen sudah sesuai dengan yang diharapkan

atau belum, sehingga nantinya instrumen dilakukan penilaian layak atau tidaknya.

2. Data Hasil Uji Coba Lapangan awal

Kegiatan uji coba lapangan awal dilakukan di SMP Negeri 5 Yogyakarta dikelas VIII H. Kegiatan uji coba lapangan awal hanya dilakukan pada skala yang lebih kecil yaitu 28 siswa. Uji coba lapangan awal dilakukan pada hari selasa 15 Maret 2019 pada jam ke- 7&8. Adapun jadwal uji coba lapangan terdapat pada lampiran 2 halaman 124. Setelah dilakukan uji coba lapangan awal diperoleh informasi waktu dalam pengerjaan tes HOTS yaitu kurang lebih 90 menit. Analisis yang digunakan untuk mendapatkan informasi mengenai setiap kualitas butir soal menggunakan analisis AnBuso versi 8.0. Berdasarkan data hasil uji coba lapangan awal tahap berikutnya adalah melakukan analisis setiap butir soal baik itu soal pilihan ganda dan uraian.

Semua data analisis diperoleh berdasarkan dari data empirik uji coba lapangan awal. Untuk data hasil analisis secara lengkap dapat dilihat pada lampiran 13 halaman 171. Kualitas instrumen penilaian HOTS dilihat berdasarkan karakteristik setiap butir soal, yang mengacu pada tingkat kesukaran butir soal, daya beda butir soal, dan keefektifan distraktor/ pengecoh. Berikut disajikan data ringkasan hasil interpretasi analisis instrumen penilaian HOTS uji coba lapangan awal.

a. Hasil Interpretasi Analisis Instrumen Penilaian HOTS Pilihan Ganda

1) Tingkat Kesukaran Butir Soal

Berdasarkan dari hasil analisis butir soal menggunakan analisis AnBuso Versi 8.0 dapat dilihat pada tabel 17.

Tabel. 17 Tingkat Kesukaran Produk Soal HOTS Pilihan Ganda

Kategori	Nomor Butir Soal	Jumlah	Persentase
$P < 0,3$ (Sukar)	3, 7, 9, 11, 14, 15, 20, 24	8	30,76%
$0,3 \leq p \leq 0,7$ (Sedang)	1, 2, 4, 5, 6, 8, 10, 12, 13, 16, 17, 18,19, 21, 22, 23, 25, 26	18	69,23%
$p > 0,7$ (Mudah)	-	0	0

Berdasarkan dari tabel 17 terlihat bahwa tingkat kesukaran butir soal belum memiliki proporsi soal yang seimbang, dimana jumlah soal kategori sedang memiliki persentase lebih tinggi yaitu berjumlah 18 soal (69,23%) pada butir soal nomor 1, 2, 4, 5, 6, 8, 10, 12, 13, 16, 17, 18,19, 21, 22, 23, 25, 26. Kategori sukar lebih rendah berjumlah 8 soal (30,76%) yaitu pada butir soal nomor 3, 7, 9, 11, 14, 15, 20, 24 dan untuk kategori mudah tidak ada sama sekali. Jika mengacu pada pendapatnya Sudjana (2017: 135) bahwa ada beberapa dasar pertimbangan dalam menentukan proporsi jumlah soal kategori mudah, sedang, dan sukar. Pertimbangan pertama adalah adanya keseimbangan, yakni jumlah soal sama untuk ketiga kategori tersebut.

Artinya soal mudah, sedang, dan sukar jumlahnya seimbang. Pertimbangan kedua proporsi jumlah soal untuk ketiga kategori tersebut didasarkan atas kurva normal. Artinya, sebagian besar soal berada dalam kategori sedang, sebagian lagi termasuk ke dalam kategori mudah dan sukar dengan proporsi yang seimbang. Perbandingan antara soal mudah, sedang, dan sukar bisa dibuat 3-4-3. Artinya 30% soal kategori mudah, 40% soal kategori sedang, dan 30% lagi soal kategori sukar. Selain itu juga bisa dibuat 3-5-2 artinya 30% soal kategori mudah, 50% soal kategori sedang, dan 20% soal kategori sukar.

2) Daya Pembeda Butir Soal

Daya pembeda butir soal dapat diketahui ketika sudah dilakukan analisis AnBuso Versi 8.0 dan berikut adalah data hasil analisis pada tabel 18.

Tabel. 18 Daya Pembeda Butir Soal HOTS Pilihan Ganda

Koefisien	Nomor Butir Soal	Jumlah	Persentase
>0,3 (Baik)	1, 2, 4, 5, 6, 7, 8, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 21, 22, 23, 25, 26	22	84,61%
0,2-0,3 (Cukup Baik)	20, 24	2	7,6%
< 0,2 (Tidak Baik)	3, 9	2	7,6%

Berdasarkan tabel 18 menunjukkan bahwa tingkat daya pembeda butir soal berada dalam kategori baik sejumlah 22 soal (84,61%) yaitu pada nomor soal 1, 2, 4, 5, 6, 7, 8, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 21, 22, 23, 25, 26. Kategori cukup baik sejumlah 2 soal

(7,6%) yaitu pada nomor soal 20, 24. Kategori tidak baik sejumlah 2 soal (7,6%) yaitu pada nomor 3, 9 artinya yang demikian harus ada beberapa soal yang perlu direvisi lebih lanjut.

3) Keefektifan Distraktor/ Pengecoh

Secara umum distraktor/ pengecoh dikatakan baik apabila dipilih oleh beberapa peserta tes atau minimal dipilih oleh 5% peserta tes (Basuki dan Hariyanto, 2014: 144). Berikut ini data keefektifan distraktor/pengecoh butir soal dapat dilihat pada tabel 19.

Tabel. 19 Keefektifan Distraktor/ Pengecoh Soal HOTS Pilihan Ganda

Kategori	Nomor Butir Soal	Jumlah	Persentase
Baik	1, 2, 4, 5, 6, 8, 10, 12, 13, 16, 17, 18, 22, 25	14	53,84%
Cukup Baik	7, 11, 14, 15, 20, 24	6	23,07%
Tidak Baik	3, 9, 19, 21, 23, 26	6	23,07%

Berdasarkan dari tabel 19 rata-rata distraktor dapat dikategorikan kedalam 3 jenis yaitu kategori baik berjumlah 14 soal (53,84%) yaitu pada nomor 1, 2, 4, 5, 6, 8, 10, 12, 13, 16, 17, 18, 22, 25. Kategori cukup baik berjumlah 6 soal (23,07%) yaitu pada nomor 7, 11, 14, 15, 20, 24. Kategori tidak baik berjumlah 6 soal (23,07%) yaitu nomor 3, 9, 19, 21, 23, 26 sehingga harus dilakukan revisi untuk nomor-nomor soal tersebut karena keberfungsian pengecoh kurang baik.

Hasil analisis karakteristik butir soal di atas, terlihat bahwa keseluruhan soal dikatakan baik. Dilihat dari tingkat kesukaran, daya pembeda butir soal, dan keefektifan distraktor/ pengecoh. Namun

untuk beberapa perlu ada revisi atau jika tidak memungkinkan direvisi maka diganti, karena daya pembeda butir soal dan keefektifan distraktor/ pengecoh kurang baik. Untuk butir soal yang kategori baik dapat digunakan sebagai produk utama instrumen penilaian HOTS dan nantinya juga akan digabungkan dengan beberapa soal yang sudah direvisi.

b. Hasil Interpretasi Analisis Instrumen Penilaian HOTS Uraian

1) Tingkat Kesukaran Butir Soal

Berdasarkan dari hasil analisis butir soal menggunakan analisis AnBuso Versi 8.0 dapat dilihat pada tabel 20.

Tabel. 20 Tingkat Kesukaran Produk Soal HOTS Uraian

Kategori	Nomor Butir Soal	Jumlah	Persentase
$P < 0,3$ (Sukar)	-	0	-
$0,3 \leq p \leq 0,7$ (Sedang)	3, 4	2	50%
$p > 0,7$ (Mudah)	1, 2	2	50%

Berdasarkan dari tabel 20 terlihat bahwa tingkat kesukaran butir soal belum seimbang. Hal ini terlihat dari hasil analisis yang dimana untuk soal kategori sukar tidak ada sama sekali. Kategori sedang berjumlah 2 soal (50%) yaitu pada butir soal nomor 3, 4. Untuk kategori mudah berjumlah 2 soal (50%) yaitu pada butir soal nomor 1, 2. Jika mengacu pada teori nya Sudjana yang sudah dijelaskan sebelumnya maka ketiga kategori soal tersebut harusnya memiliki persentase soal masing-masing agar seimbang. Jadi untuk dua soal uraian yang memiliki kategori mudah yaitu nomor 1, 2

perlu untuk dilihat kembali apakah memang soal sudah sesuai dengan indikator soal HOTS atau belum.

2) Daya Pembeda Butir Soal

Berdasarkan dari data analisis daya pembeda butir soal dapat dilihat pada tabel 21.

Tabel. 21 Daya Pembeda Butir Soal HOTS Uraian

Koefisien	Nomor Butir Soal	Jumlah	Persentase
>0,3 (Baik)	1, 2, 3	3	75%
0,2-0,3 (Cukup Baik)	4	1	25%
< 0,2 (Tidak Baik)	-	0	-

Berdasarkan tabel 21 menunjukkan bahwa tingkat daya pembeda butir soal berada dalam kategori baik dan cukup baik. Kategori baik sejumlah 3 soal (75%) yaitu pada nomor soal 1, 2, 3. Kategori cukup baik sejumlah 1 soal (25%) yaitu pada nomor soal 4.

3. Data Hasil Uji Coba Lapangan

Uji coba lapangan dilakukan beberapa hari setelah uji coba lapangan awal. Uji coba lapangan awal dilakukan di kelas 8B, 8C, 8D, 8E, 8F, 8G, dan 8I jumlah keseluruhan mencapai 200 siswa. Untuk kegiatan uji coba lapangan dapat dilihat pada jadwal yang dilampirkan pada lampiran 2 halaman 124. Data uji coba lapangan diperoleh berdasarkan data empirik dilapangan untuk melihat kualitas produk yang dikembangkan yaitu instrumen penilaian HOTS. Kualitas instrumen penilaian HOTS dianalisis menggunakan aplikasi AnBuso versi 8.0 untuk jenis soal pilihan ganda dan

uraian. Analisis dilakukan pada setiap butir soal dengan melihat nilai tingkat kesukaran, daya beda, dan distraktor/ pengecoh. Hasil analisis butir soal secara keseluruhan disajikan pada lampiran 14 halaman 173. Berikut ringkasan hasil interpretasi analisis instrumen penilaian HOTS uji coba lapangan.

a. Hasil Interpretasi Analisis Instrumen Penilaian HOTS Pilihan Ganda

1) Tingkat Kesukaran Butir Soal

Secara keseluruhan hasil analisis berdasarkan data yang diperoleh dilapangan mengenai tingkat kesukaran butir soal disajikan pada tabel 22.

Tabel. 22 Tingkat Kesukaran Produk Soal HOTS Pilihan Ganda

Kategori	Nomor Butir Soal	Jumlah	Persentase
$P < 0,3$ (Sukar)	1, 2, 8, 11, 12, 16, 21, 23, 24, 26	10	38,46%
$0,3 \leq p \leq 0,7$ (Sedang)	3, 4, 5, 6, 7, 9, 10, 13, 14, 15, 17, 18, 19, 20, 22, 25	16	61,53%
$p > 0,7$ (Mudah)	-	0	-

Berdasarkan dari tabel 22 terlihat bahwa tingkat kesukaran butir soal belum seimbang. Hal ini terlihat dari hasil analisis yang dimana untuk soal kategori sukar persentase nya lebih sedikit dari soal kategori sedang. Untuk kategori soal mudah persentasenya tidak ada sama sekali. Kategori sukar sejumlah 10 soal (38,46%) yaitu pada nomor soal 1, 2, 8, 11, 12, 16, 21, 23, 24, 26. Kategori sedang sejumlah 16 soal (61,53%) yaitu pada nomor soal 3, 4, 5, 6, 7, 9, 10, 13, 14, 15, 17, 18, 19, 20, 22, 25. Jika mengacu pada teori

nya Sudjana yang sudah dijelaskan sebelumnya maka ketiga kategori soal tersebut harusnya memiliki persentase soal masing-masing agar seimbang.

2) Daya Pembeda Butir Soal

Daya pembeda butir soal sangat penting untuk melihat kualitas butir soal apakah sudah baik, cukup baik, atau tidak baik. Analisis daya pembeda butir soal diperoleh dari analisis AnBuso Versi 8.0 dan berikut adalah data hasil analisis pada tabel 23.

Tabel. 23 Daya Pembeda Butir Soal HOTS Pilihan Ganda

Koefisien	Nomor Butir Soal	Jumlah	Persentase
>0,3 (Baik)	1, 2, 3, 5, 6, 7, 8, 9, 10, 12, 13, 14, 15, 17, 18, 19, 20, 21, 22, 25	20	76,93%
0,2-0,3 (Cukup Baik)	11, 16, 24, 26	4	15,38%
< 0,2 (Tidak Baik)	4, 23	2	7,6%

Berdasarkan tabel 23 menunjukkan bahwa rata-rata daya beda butir soal berada pada kategori baik yaitu 20 soal (76,93%) yaitu pada nomor soal 1, 2, 3, 5, 6, 7, 8, 9, 10, 12, 13, 14, 15, 17, 18, 19, 20, 21, 22, 25. Kategori cukup baik sejumlah 4 soal (15,38%) yaitu soal nomor 11, 16, 24, 26. Kategori tidak baik sejumlah 2 soal (7,6%) yaitu soal nomor 4, 23 hal tersebut maka soal harus direvisi kembali agar mendapatkan kualitas butir soal yang baik yang nantinya dapat dijadikan sebagai produk utama instrumen penilaian HOTS.

3) Keefektifan Distraktor/ Pengecoh

Secara umum distraktor/ pengecoh dikatakan baik apabila dipilih oleh beberapa peserta tes atau minimal dipilih oleh 5% peserta tes (Basuki dan Hariyanto, 2014: 144). Berikut ini data keefektifan distraktor/pengecoh butir soal dapat dilihat pada tabel 24.

Tabel. 24 Keefektifan Distraktor/ Pengecoh Soal HOTS Pilihan Ganda

Kategori	Nomor Butir Soal	Jumlah	Persentase
Baik	3, 5, 6, 7, 9, 10, 13, 14, 15, 17, 18, 19, 20, 22, 25	15	57,69%
Cukup Baik	1, 2, 8, 11, 12, 16, 21, 24, 26	9	34,61%
Tidak Baik	4, 23	2	7,6%

Berdasarkan dari tabel 24 rata-rata distraktor dapat dikategorikan kedalam 3 jenis yaitu kategori baik berjumlah 15 soal (57,69%) yaitu pada nomor 3, 5, 6, 7, 9, 10, 13, 14, 15, 17, 18, 19, 20, 22, 25. Kategori cukup baik berjumlah 9 soal (34,61%) yaitu pada nomor 1, 2, 8, 11, 12, 16, 21, 24, 26. Kategori tidak baik berjumlah 2 soal (7,6%) yaitu nomor 4, 23 sehingga harus dilakukan revisi untuk kedua nomor soal tersebut.

b. Hasil Interpretasi Analisis Instrumen Penilaian HOTS Uraian

1) Tingkat Kesukaran Butir Soal

Berdasarkan dari hasil analisis butir soal menggunakan analisis AnBuso Versi 8.0 dapat dilihat pada tabel 25.

Tabel. 25 Tingkat Kesukaran Produk Soal HOTS Uraian

Kategori	Nomor Butir Soal	Jumlah	Persentase
$P < 0,3$ (Sukar)	-	0	-
$0,3 \leq p \leq 0,7$ (Sedang)	1, 3, 4	3	75%
$p > 0,7$ (Mudah)	2	1	25%

Berdasarkan dari tabel 25 terlihat bahwa tingkat kesukaran butir soal belum seimbang. Hal ini terlihat dari hasil analisis yang dimana untuk soal kategori sukar persentasenya tidak ada sama sekali, kategori sedang lebih tinggi dari kategori soal kategori mudah. Tingkat kesukaran berkisar hanya pada kategori sedang dan mudah. Kategori sedang berjumlah 3 soal (75%) yaitu pada butir soal nomor 1, 3, 4. Untuk kategori mudah berjumlah 1 soal (25%) yaitu pada butir soal nomor 2. Jika mengacu pada teorinya Sudjana yang sudah dijelaskan sebelumnya maka ketiga kategori soal tersebut harusnya memiliki persentase soal masing-masing agar seimbang. Jadi untuk soal uraian yang memiliki kategori mudah yaitu nomor 2 perlu untuk dilihat kembali apakah memang soal sudah sesuai dengan indikator soal HOTS atau belum.

2) Daya Pembeda Butir Soal

Berdasarkan dari data analisis daya pembeda butir soal dapat dilihat pada tabel 26.

Tabel. 26 Daya Pembeda Butir Soal HOTS Uraian

Koefisien	Nomor Butir Soal	Jumlah	Persentase
>0,3 (Baik)	1, 2, 3, 4	4	100%
0,2-0,3 (Cukup Baik)	-	0	-
< 0,2 (Tidak Baik)	-	0	-

Berdasarkan tabel 26 menunjukkan bahwa butir soal uraian memiliki tingkat daya beda yang berada dalam kategori baik semua.

C. Kemampuan *Higher Order Thinking Skills* Siswa SMP Negeri 5 Yogyakarta Pada Mata Pelajaran IPS

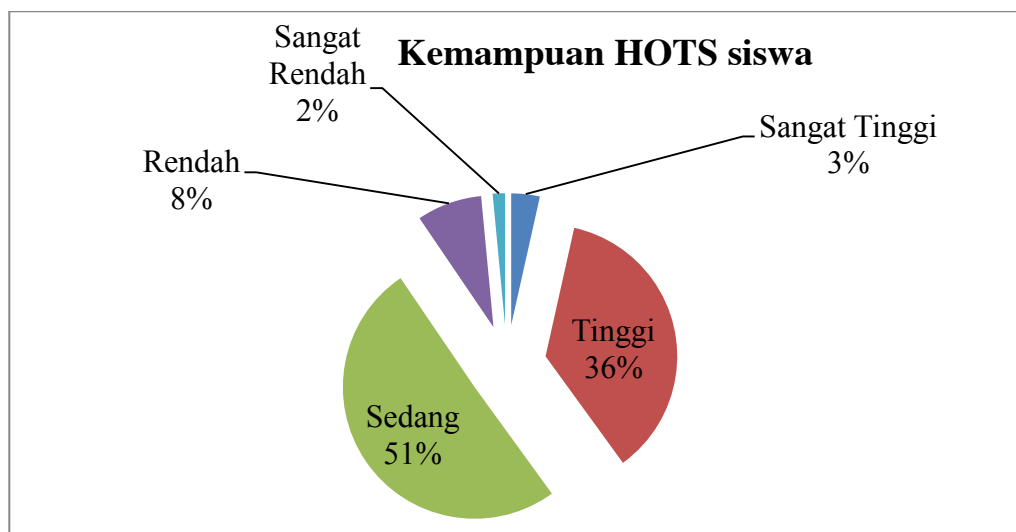
Berdasarkan dari hasil yang telah dilakukan diperoleh data kemampuan HOTS siswa SMP Negeri 5 Yogyakarta, pada mata pelajaran IPS. Tingkat kemampuan HOTS siswa dapat dikategorikan kedalam lima tingkatan. Berikut ini disajikan data hasil belajar siswa berdasarkan kelas interval.

Tabel. 27 Interval Kemampuan HOTS Siswa SMP N 5 Yogyakarta

No	Interval Kemampuan	Kategori	Jumlah	Persentase
1.	85-100%	Sangat Tinggi	7	3,5%
2.	75-84%	Tinggi	73	36,5%
3.	60-74%	Sedang	101	50,5%
4.	40-59%	Rendah	16	8%
5.	0-39%	Sangat Rendah	3	1,5%

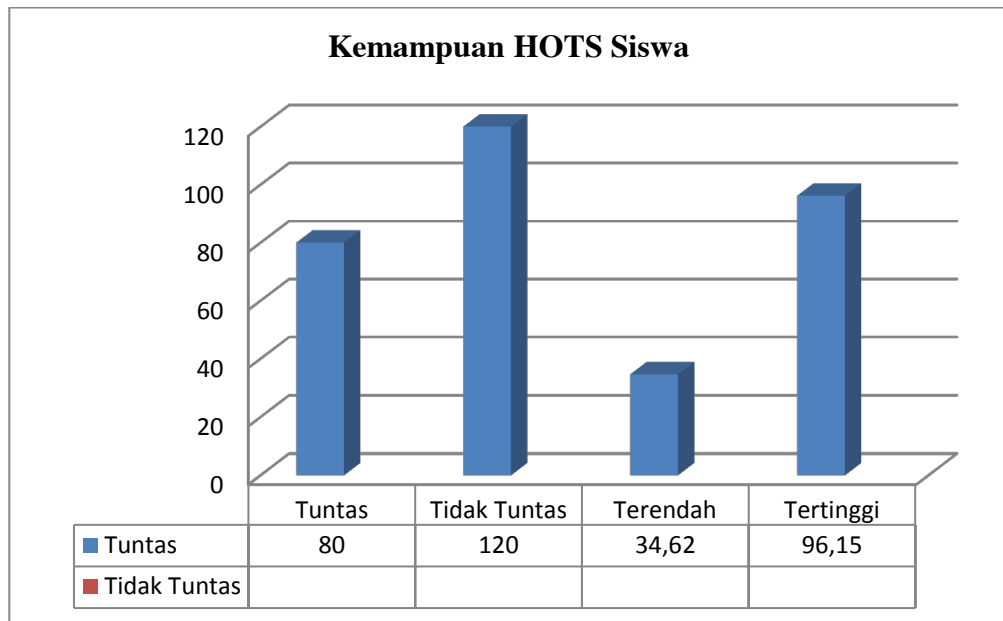
Berdasarkan dari tabel di atas dapat dilihat bahwa kemampuan HOTS siswa pada mata pelajaran IPS berbeda-beda. Dari yang sangat tinggi memiliki persentase (3,5%), kategori tinggi (36,5%), kategori sedang (50,5%), kategori rendah (8%), dan kategori sangat rendah (1,5%).

Berdasarkan dari data tersebut dapat disimpulkan bahwa rata-rata kemampuan *higher order thinking skills* siswa SMP Negeri 5 Yogyakarta berada pada posisi sedang sampai tinggi. Sedangkan untuk kategori rendah dan sangat rendah persentasenya jauh lebih sedikit. Hasil nilai siswa secara keseluruhan dapat dilihat pada lampiran 17 halaman 186. Berikut ini disajikan gambar hasil kemampuan HOTS siswa SMP Negeri 5 Yogyakarta.



Gambar. 7 Kemampuan HOTS Siswa SMP Negeri 5 Yogyakarta

Adapun secara keseluruhan ketuntasan belajar siswa dalam mengerjakan tes HOTS dapat dilihat pada gambar di bawah ini.



Gambar. 8 Hasil Belajar Siswa

Berdasarkan dari gambar tersebut dapat dilihat bahwa jumlah siswa yang mengikuti tes penilaian HOTS berjumlah 200 siswa. Siswa yang tuntas adalah 80 siswa dan siswa yang tidak tuntas adalah 120. Sedangkan nilai tertinggi hasil tes penilaian HOTS siswa mencapai 96,15 sementara nilai terendahnya adalah 34,62. Berdasarkan kedua gambar tersebut dapat dilihat tingkat kemampuan HOTS siswa SMP Negeri 5 Yogyakarta pada mata pelajaran IPS berada pada kategori baik yaitu dikategori sedang dan tinggi.

Dilihat dari hasil analisis kuesioner juga menunjukkan untuk tiap-tiap indikator adalah (1) kemampuan menganalisis lebih dominan mempengaruhi HOTS siswa pada mata pelajaran IPS. Hal ini dapat dilihat dari nilai *eigenvalue* dan variannya yang paling tinggi yaitu 5,943 dan 39,623%. (2) kemampuan mengevaluasi menunjukkan nilai *eigenvalue* dan varian yaitu 0,446 dan 2,970%. (3) kemampuan mencipta menunjukkan nilai *eigenvalue*

dan varian 0,358 dan 2,388%. Data hasil analisis kuesioner disajikan pada lampiran 12 halaman 160.

Berdasarkan dari penelitian sebelumnya oleh Gusmira Wita (2018) juga menunjukkan bahwa kemampuan menganalisis dominan kedua setelah kemampuan menyelesaikan masalah. Ini terlihat dari nilai *eigenvalue* 2,326 dan varian 8,022%. Terkait dengan kemampuan menganalisis (Anderson & Krathwohl, 2012: 120-133) mengungkapkan bahwa kemampuan menganalisis merupakan kemampuan menganalisis informasi yang dikaitkan ke dalam bagian-bagian sehingga dapat mengenali pola dan hubungannya yang kemudian dapat diketahui sebab akibat dari sebuah permasalahan.

Kemampuan mengevaluasi menunjukkan nilai *eigenvalue* 0,446 dan varian 2,970%. Mengevaluasi adalah kemampuan dalam pemberian penilaian terhadap solusi, gagasan, dan metodologi. Proses mengevaluasi memiliki kriteria atau standar tersendiri yang dianggap cocok, sehingga dapat diperoleh nilai yang ada efektivitas dan manfaatnya (Anderson & Krathwohl, 2012: 120-133). Kaitannya kemampuan menganalisis dengan kemampuan *Higher order thinking skills* merupakan kemampuan yang tidak hanya menghapalkan materi dan menerapkan materi saja melainkan lebih kepada kegiatan bagaimana mengevaluasi suatu materi (Jones, 2016: 262).

Kemampuan mencipta menunjukkan nilai *eigenvalue* 0,358 dan varian 2,388%. Terkait dengan kemampuan mencipta (Anderson & Krathwohl, 2012: 120-133) mencipta didefinisikan sebagai menggeneralisasikan ide baru, produk atau cara pandang yang baru dari sesuatu kejadian. Mencipta disini

diartikan sebagai merancang sesuatu guna untuk menyelesaikan masalah. Mencipta bisa berupa wujud ide atau gagasan dari siswa mengenai rencana aksi. Kegiatan mencipta umumnya berhubungan dengan pengalaman belajar siswa yang sebelumnya.

Sebuah proses pembelajaran yang dilaksanakan sangat berhubungan erat dengan *assesment* pembelajaran yang dilaksanakan. Keberhasilan sebuah proses pembelajaran dapat dilihat dari *assesment* yang dilakukan (Marzano, 1994: 11). Proses penilaian juga dapat dikatakan sebagai indikator dari proses pembelajaran itu sendiri, apakah proses pembelajaran itu berhasil atau tidak. Dari hasil tes penilaian berbasis HOTS mata pelajaran IPS dapat memberikan gambaran umum tentang proses pembelajaran IPS di kelas.

Berdasarkan dari hasil yang diperoleh melalui tes penilaian berbasis HOTS yang diberikan kepada siswa pada pembelajaran IPS, dapat disimpulkan bahwa sebenarnya mengajarkan HOTS kepada siswa bukanlah sesuatu hal yang sulit. Hal ini tergantung dari cara mengajar guru, sebelum siswanya dituntut untuk memiliki kemampuan berpikir tingkat tinggi terlebih dahulu memberikan pemahaman dan pengetahuan tentang HOTS kepada guru itu sendiri. Guru harus sengaja merancang pelajaran atau program pembelajaran untuk menguatkan strategi berpikir (Kauchak & Eggen, 1998 Darmer, 1995) (Sani, 2019: 80). Strategi pembelajaran harus sesuai dengan karakteristik siswa dan materi untuk mengembangkan kemampuan beripikir HOTS siswa.

Menurut Nugroho (2019: 67) membiasakan HOTS kepada siswa tidak bisa dilakukan secara tiba-tiba dan instan. Membiasakan HOTS membutuhkan strategi holistik dari para guru. Guru tidak dapat menagih siswa dengan pengukuran dan *assesmen* bertipe HOTS diakhir pembelajaran tanpa melakukan pembelajaran HOTS terlebih dulu. HOTS harus didesain dengan matang sesuai dengan konteks siswa dan materi ajar. Miri, et al (2007: 363) mengusulkan tiga strategi pembelajaran yang dapat digunakan untuk mengembangkan HOTS: 1) proses pembelajaran berbasis masalah yang kontekstual yang dapat disajikan di dalam kelas. 2) diskusi terbuka, 3) pembelajaran yang diarahkan pada kegiatan penyelidikan dan penelitian.

Guru dianjurkan untuk menerapkan beberapa prinsip dalam pembelajaran berbasis HOTS, antara lain: (1) memberikan tugas yang sesuai dengan tujuan materi dan kemampuan siswa, (2) menumbuhkan rasa ingin tahu kepada siswa, (3) memberikan tugas kepada siswa agar siswa aktif berpikir, (4) membahas persoalan yang kontekstual, (5) mengembangkan imajinasi kepada siswa, (6) memberi kesempatan kepada siswa untuk berpikir divergen (lateral) dan elaborasi, (7) memberikan kesempatan kepada siswa untuk memilih topik pelajaran yang akan dipelajari, (8) tidak mudah menyalahkan siswa, (9) siswa diberi kesempatan untuk bereksperimen, (10) memberikan *reward* kepada siswa, (11) siswa dilatih untuk bertanya, menyusun rumusan masalah, dan menyelesaikan masalah, (12) siswa dilatih untuk berpikir kritis, menganalisis, dan mengevaluasi, (13) siswa dilatih untuk dapat membuat keputusan (Sani,2019: 71-72)

Pendapat lain juga disampaikan oleh Zohar & Nemet (2002: 39) mengemukakan bahwa aktifitas HOTS dapat berupa aktifitas mengeluarkan argumentasi/pendapat, melakukan perbandingan, memberikan kesepakatan dengan pendapat yang berbeda, melakukan pemecahan masalah, mengambil keputusan, serta mengidentifikasi asumsi yang tersembunyi. Dengan demikian peserta didik harus sering dilatih untuk dapat mengembangkan kemampuan berpikir tingkat tinggi, dengan cara menerapkan aktifitas belajar yang dapat mendorong kemampuan HOTS dan lebih sering dalam berlatih mengerjakan soal jenis-jenis HOTS. Aktifitas-aktifitas belajar yang dapat mendorong kemampuan HOTS tersebut adalah diskusi, pembelajaran berbasis masalah yang kontekstual, dan melakukan perbandingan.

Aktifitas-aktifitas belajar yang demikian yang dapat melatih siswa untuk aktif bertanya, berpendapat, berani mengambil keputusan, dan kemampuan pemecahan masalah. Manfaat dari proses pembelajaran berbasis HOTS adalah dapat membuat siswa antusias dalam belajar, memotivasi siswa, dan membentuk mental siswa agar tidak mudah menyerah. Jika siswa dihadapkan dengan soal-soal yang menuntut berpikir tingkat tinggi maka siswa akan merasa tertantang untuk dapat menyelesaikannya. Dengan demikian prestasi siswa juga dapat meningkat dan mampu menyelesaikan masalah yang kontekstual sehingga kesadaran dan kepekaan sosial didalam diri siswa akan terbentuk. Selain itu juga siswa akan menjadi warga masyarakat yang berguna dan bermanfaat dilingkungannya.

D. Revisi Produk

Revisi produk dilakukan untuk menyempurnakan butir soal HOTS agar memiliki kualitas yang baik. Selain itu revisi produk juga untuk memenuhi setiap butir soal agar soal bisa dikatakan valid dan reliabel. Kegiatan revisi produk didasarkan pada hasil validasi dan uji coba lapangan awal.

1. Revisi Produk Setelah Validasi



Validasi dilakukan oleh tiga dosen UNY untuk menilai dan memberi komentar serta saran pada produk yang sedang dikembangkan. Berdasarkan hasil validasi instrumen dari dosen terdapat beberapa masukan yang pertama untuk soal nomor 12 rumusan soal masih belum jelas dan jawaban sudah terjawab didalam soal. Kedua untuk soal nomor 16 stetmen soal terlalu berpuisi karena antara stetmen dan pertanyaan soal tidak berhubungan. Ketiga untuk soal nomor 23 yang dimana dalam soal tersebut daya beda pernyataan tidak jelas. Keempat untuk soal nomor 24 soal gambar tidak berfungsi dengan jelas, karena dapat menimbulkan persepsi ganda pada siswa.

Hasil validasi untuk soal uraian ada beberapa yang perlu direvisi yaitu soal nomor 4. Revisi dilakukan dengan mengganti jenis pertanyaan yang semula jenis pertanyaan “sebutkan” dan diganti dengan jenis pertanyaan “bagaimana” karena ini lebih dapat mengajak siswa untuk berpikir kritis. Beberapa butir soal yang lainnya yang perlu direvisi adalah terkait dengan kaidah penulisan soal dan SPOK yang belum sesuai. Masukan dari validator yang lainnya adalah terkait dengan kesesuaian soal terhadap

indikator HOTS dan rata-rata pengecoh yang masih kurang. Dari beberapa komentar dan masukan dari dosen validator tersebut perlu untuk diperhatikan sebagai bahan revisi. Berikut ini disajikan rangkuman hasil revisi berdasarkan masukan dari validator pada tabel 28.

Tabel. 28 Hasil Revisi Berdasarkan Validator

Soal Pilihan Ganda		
No	Sebelum Direvisi	Sesudah Direvisi
12	<p>Pada tahun 2017 pemerintah Indonesia mengimpor gula konsumsi sekitar 1,3 juta ton dari berbagai negara, salah satunya Thailand. Impor gula untuk memenuhi kebutuhan industri sebanyak 3,5 juta ton. Alasan yang mendasari Indonesia masih melakukan impor gula dari Thailand adalah ...</p> <ol style="list-style-type: none"> Adanya persamaan selera antara masyarakat Indonesia dan Thailand Indonesia tidak dapat menghasilkan tebu dengan kualitas baik Menghindari kelangkaan persediaan gula dalam negeri Meningkatkan pendapatan negara untuk satu tahun 	<p>Kegiatan ekspor & impor saat ini semakin meningkat karena adanya kemajuan teknologi. Meningkatnya arus kegiatan ekspor & impor menyebabkan produk-produk dijual dalam pasar bebas. Berdasarkan dari kegiatan tersebut pemerintah mengeluarkan kebijakan kuota, yang dimaksud kebijakan kouta adalah ...</p> <ol style="list-style-type: none"> membuat aturan untuk meniadakan peredaran barang tertentu dari luar negeri mengurangi jumlah barang dan/ atau jasa yang didatangkan dari negara lain mengenakan biaya tinggi bagi eksportir dari luar negeri memberi keringanan biaya bagi pengusaha/ eksportir Indonesia
16	<p>Indonesia sebagai negara yang memiliki kekayaan laut yang melimpah juga memiliki sektor agrikultural yang cukup baik, sebagai penghasil pangan dalam pemenuhan kebutuhan masyarakat. Salah satu peran pemerintah dalam penguatan ekonomi agrikultural di Indonesia ...</p>	<p>Indonesia sebagai negara yang memiliki sektor agrikultur yang cukup baik sebagai penghasil dalam pemenuhan kebutuhan pangan masyarakat. Menurut pendapat anda bagaimana strategi yang tepat untuk mengembangkan agrikultur di Indonesia ...</p> <ol style="list-style-type: none"> melestarikan cara pertanian tradisional

	<p>a. Memberi subsidi pupuk serta benih</p> <p>b. Membuka hutan untuk lahan pertanian</p> <p>c. Membersihkan lahan pertanian secara gratis</p> <p>d. Melakukan impor beras berkualitas dari negara lain</p>	<p>b. mendatangkan teknologi yang canggih dari negara lain</p> <p>c. menerapkan sistem pertanian secara ekofarming yang ramah lingkungan</p> <p>d. membuka lahan pertanian dan memberikan kepada masyarakat secara gratis</p>
23	<p>Perhatikan pernyataan berikut!</p> <ol style="list-style-type: none"> 1) Menyerap tenaga kerja di masyarakat sekitar 2) Penyedia bahan pangan 3) Satu-satunya sumber pendapatan negara 4) Penyedia bahan baku industri 5) Penyeimbang ekosistem dengan adanya banyak persawahan <p>Dari data di atas yang menunjukkan manfaat dari sektor pertanian ditunjukkan oleh angka ...</p> <ol style="list-style-type: none"> a. 1), 2), dan 3) b. 1), 2), dan 4) c. 2), 3), dan 5) d. 3), 4), dan 5) 	<p>Perhatikan pernyataan berikut!</p> <ol style="list-style-type: none"> 1) Menyerap tenaga kerja di masyarakat sekitar 2) Penyedia bahan pangan 3) Penyumbang ekspor terbesar 4) Penyedia bahan baku industri 5) Penyeimbang ekosistem dengan adanya banyak persawahan <p>Dari data di atas, manfaat dari sektor pertanian ditunjukkan oleh angka ...</p> <ol style="list-style-type: none"> a. 1), 2), dan 4) b. 1), 2), dan 3) c. 2), 3), dan 5) d. 3), 4), dan 5)
24	<p>Perhatikan gambar berikut!</p>  <p>Berdasarkan gambar tersebut terlihat suatu permasalahan di Indonesia dalam ...</p> <ol style="list-style-type: none"> a. Menyediakan lahan permukiman bagi masyarakat b. Pendistribusian pendapatan nasional 	<p>Perhatikan gambar berikut!</p>  <p>Gambar tersebut menunjukkan permasalahan yang terjadi di Indonesia terutama di kota-kota besar. Menurut pendapat anda permasalahan tersebut muncul karena ...</p> <ol style="list-style-type: none"> a. pemerintah kurang menyediakan lahan

	<ul style="list-style-type: none"> c. Bantuan bahan bangunan tidak merata d. Perencanaan tata wilayah dan kota 	<ul style="list-style-type: none"> permukiman bagi masyarakat di kota b. perencanaan dan tata ruang kota yang belum dilakukan secara maksimal c. pemberian bantuan kepada masyarakat miskin yang tidak tepat sasaran d. pendistribusian pendapatan nasional yang tidak merata
Soal Uraian		
	Sebelum Direvisi	Sesudah Direvisi
4	<p>Redistribusi (pendistribusian kembali) pendapatan adalah pendistribusian kembali pendapatan masyarakat kelompok kaya kepada masyarakat kelompok miskin baik berasal dari pajak ataupun pungutan-pungutan lain. Sebutkan program dari pemerintah dalam rangka redistribusi pendapatan!</p>	<p>Redistribusi (pendistribusian kembali) pendapatan adalah pendistribusian kembali pendapatan masyarakat kelompok kaya kepada masyarakat kelompok miskin, baik berasal dari pajak ataupun pungutan-pungutan lain. Bagaimana pendapat anda mengenai program redistribusi pendapatan?</p>

Secara keseluruhan soal-soal yang lainnya sudah sesuai dengan indikator, namun ada sebagian soal yang harus diperbaiki terkait dengan rumusan kalimat agar sesuai dengan kaidah penulisan soal yang benar. Kesalahan yang banyak terjadi pada semua soal yaitu kaidah penulisan yang masih salah dan SPOK yang masih belum tepat. Kesalahan terjadi pada pilihan jawaban yang dimana di awal kalimat menggunakan huruf kapital, jadi dilakukan revisi untuk semua soal. Kaidah penulisan yang masih salah juga terdapat pada soal nomor 11 yang dimana nama negara tidak menggunakan huruf kapital maka perlu untuk direvisi. Perbaikan juga dilakukan pada butir soal dan jawaban yang harusnya dibuat pada halaman yang sama. Hal ini lebih memudahkan siswa mengerjakan soal

tes. Revisi dari validator ini dimaksudkan agar instrumen penilaian mempunyai kualitas yang baik, agar selanjutnya dapat digunakan pada uji coba lapangan.

2. Revisi Produk Setelah Uji Coba

Berdasarkan dari hasil analisis dengan menggunakan AnBuso versi 8.0 dapat dilihat beberapa butir soal tidak memenuhi kriteria. Butir soal yang tidak memenuhi kriteria tersebut untuk kemudian direvisi atau diganti. Adapun soal yang tidak memenuhi kriteria tersebut adalah soal nomor 3 dan 9 karena daya beda dan pengecoh tidak baik. Bagian soal uraian yang perlu untuk direvisi karena tidak memenuhi kriteria adalah soal nomor 1 dan 2 karena distribusi kedua soal tersebut dianggap mudah. Butir soal yang dianggap mudah tersebut tidak dapat menggambarkan kemampuan tinggi dan rendah. Revisi dilakukan untuk menaikkan tingkat kesukaran soal pada kedua nomor soal tersebut, yang dimana nantinya bisa membedakan kemampuan siswa. Semua butir soal yang sudah direvisi selanjutnya digabungkan dengan soal-soal tidak direvisi untuk dapat digunakan pada uji coba lapangan dalam skala yang lebih luas. Jumlah soal yang siap untuk digunakan pada uji coba lapangan sebanyak 30 butir soal. Butir soal sebanyak 30 tersebut sudah memenuhi kriteria soal yang baik.

E. Kajian Produk Akhir

Produk akhir yang dihasilkan dari penelitian ini adalah instrumen penilaian berupa soal HOTS mata pelajaran IPS kelas VIII semester 2. Sebelum dinyatakan sebagai produk akhir terlebih dahulu instrumen melewati penilaian validitas dari dosen validator, tahap uji coba lapangan awal, hingga pada tahap uji coba lapangan skala yang lebih luas. Setelah melalui beberapa tahapan tersebut dilakukan analisis untuk melihat kualitas produk yang dikembangkan. Berdasarkan hasil analisis dapat disimpulkan bahwa produk instrumen penilaian HOTS memiliki kriteria valid dan reliabel, serta kualitas soal yang baik. Produk dikatakan valid karena telah melalui validasi dari ahli dan penilaian dari dosen validasi untuk selanjutnya dianalisis menggunakan SPSS 23.

Sedangkan produk memenuhi kriteria reliabel berdasarkan data empirik dari uji coba lapangan yang dianalisis menggunakan SPSS 23. Untuk kualitas butir soal pilihan ganda dan uraian memiliki kriteria baik berdasarkan hasil analisis menggunakan AnBuso 8.0. Selanjutnya butir soal yang siap digunakan berjumlah 30 soal, 26 untuk jenis soal pilihan ganda dan 4 soal untuk jenis soal uraian. Produk soal HOTS mata pelajaran IPS ini yang nantinya diharapkan dapat membantu guru dalam membuat soal HOTS IPS materi-materi yang lainnya. Semakin banyak guru dalam mengembangkan soal-soal HOTS mata pelajaran IPS maka kemampuan siswa akan terus berkembang.

1. Kevalidan Instrumen

Berdasarkan dari hasil analisis kevalidan instrumen penilaian HOTS yang berupa soal pilihan ganda dan uraian dapat dinyatakan valid dan layak digunakan. kevalidan tersebut diperoleh dari hasil validasi produk pengembangan yang dilakukan oleh tiga dosen ahli. Validasi oleh tiga dosen ahli pada produk instrumen penilaian HOTS memenuhi validitas logis. Ketiga validator menyatakan bahwa produk penilaian HOTS berupa soal pilhan ganda dan uraian yang dikembangkan sudah memenuhi kriteria. Hasil validasi juga dianalisis menggunakan formula Aiken's yang hasilnya dapat dilihat langsung apakah instrumen tersebut valid atau tidaknya. Hasil analisis formula Aiken's dapat dilihat pada point hasil validasi ahli.

2. Reliabilitas Instrumen

Teknik analisis data untuk estimasi reliabilitas instrumen ini menggunakan teknik estimasi konsistensi internal dengan formula *Cronbach-alpha* yang dibantu aplikasi SPSS IBM 23. Berikut ini adalah tabel kategori koefisien (Sugiyono, 2015: 257)

Tabel. 29 Kategori Koefisien Reliabilitas

No.	Interval Koefisien	Tingkat Hubungan
1.	0,00-0,199	Sangat rendah
2.	0,20-0399	Rendah
3.	0,40-0,599	Sedang
4.	0,60-0,799	Kuat
5.	0,80-1,000	Sangat kuat

Berdasarkan dari hasil analisis SPSS 23 reliabilitas untuk soal pilihan ganda dan uraian berada pada kategori yang baik jika dilihat pada tabel 29.

Nilai reliabilitas untuk soal pilihan ganda adalah 0,768 yaitu berada pada kategori kuat. Sedangkan untuk soal uraian adalah 0,804 yang berada pada kategori sangat kuat. Berikut ini disajikan data hasil analisis reliabilitas soal pilihan ganda dan uraian menggunakan SPSS 23.

Reliabilitas Soal Pilihan Ganda

Reliability Statistics		
Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
,768	1,000	27

Reliabilitas Soal Uraian

Reliability Statistics		
Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
,804	,876	5

F. Keterbatasan Penelitian

Pelaksanaan penelitian ini telah selesai dilakukan peneliti sudah mengupayakan semaksimal mungkin sesuai dengan rencana dan tujuan penelitian yang sudah ditetapkan. Namun peneliti juga mengakui adanya keterbatasan dalam pelaksanaan penelitian ini, masih banyak kelemahan-kelemahan dalam penelitian ini. *Pertama* hasil analisis data masih menggunakan analisis klasik sehingga hasil pengukuran dipengaruhi oleh responden. Untuk penelitian yang sejenis mungkin bisa menggunakan analisis modern yang dimana hasil pengukuran tidak bergantung dari responden. *Kedua* untuk proses validasi produk instrumen penilaian HOTS peneliti disini

hanya melibatkan dosen ahli yang berasal dari akademisi Pascasarjana UNY dan belum melibatkan guru ataupun teman sejawat. *Ketiga* kegiatan uji coba lapangan dilakukan pada semester 2 lebih tepatnya setelah kegiatan UTS di sekolah.

Dalam pengumpulan data tidak dapat dipungkiri bahwa responden kurang bersungguh-sungguh dalam mengerjakan, hal ini mengingat waktu uji coba lapangan yang dilaksanakan setelah kegiatan UTS selain itu juga SK dan KD nya pun materi sebelum UTS sehingga kemungkinan banyak materi yang sudah terlupakan. Untuk selanjutnya bila ada penelitian yang sejenis usahakan materi dalam pembuatan soal HOTS sejalan dengan yang ada dilapangan. *Keempat* analisis kemampuan HOTS pada mata pelajaran IPS ini hanya dilaksanakan disatu sekolah yaitu SMP Negeri 5 Yogyakarta. Untuk penelitian selanjutnya yang sejenis mungkin bisa menggunakan sampel yang lebih luas atau paling tidak dua sekolah sehingga dapat dikomparasikan kemampuan HOTS siswa pada mata pelajaran IPS.

Peneliti sangat menyadari bahwa penelitian ini jauh dari sempurna, yang disebabkan oleh beberapa hal. Peneliti hanya dapat berharap semoga hasil penelitian ini dapat menggambarkan tingkat kemampuan siswa kaitannya dalam soal-soal berbasis HOTS dan memberi gambaran proses pembelajaran IPS di sekolah. Selain itu produk yang dihasilkan yaitu instrumen penilaian HOTS dapat memberikan kontribusi pengetahuan tentang soal berbasis HOTS pada mata pelajaran IPS.