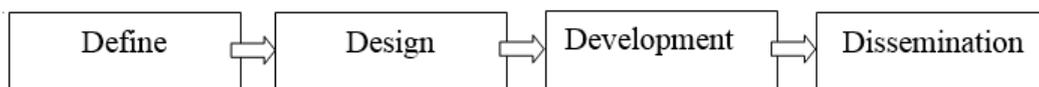


BAB III

METODE PENELITIAN

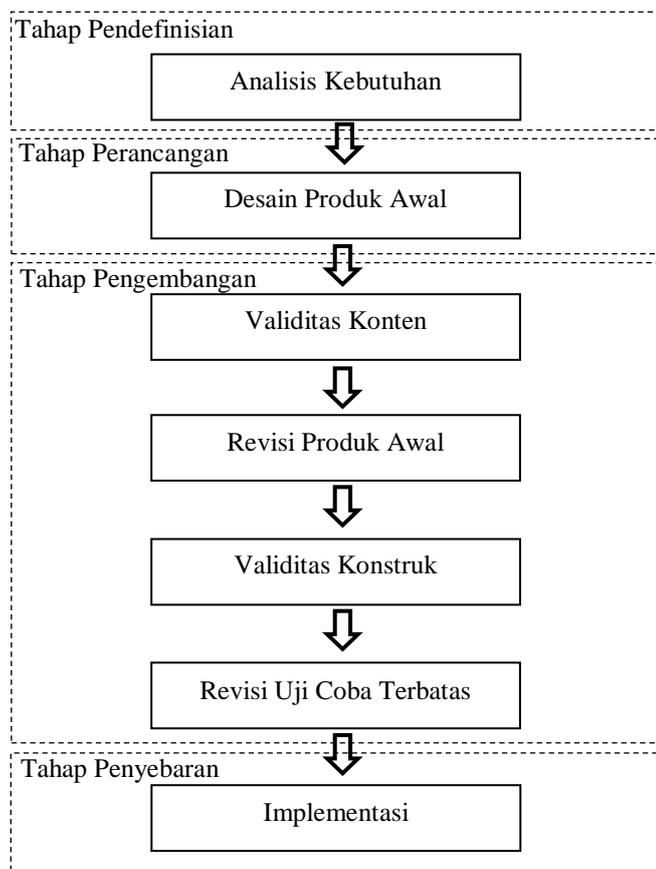
A. Desain Penelitian

Tesis ini merupakan sebuah penelitian pengembangan yang berusaha mengembangkan produk berupa asesmen kognitif aspek HOTS dengan model 4-D (*Four D*). Model 4-D dikembangkan oleh Thiagarajan dan terdiri dari empat langkah yakni pendefinisian (*define*), perancangan (*design*), pengembangan (*develope*), dan penyebaran (*dissemination*) (Thiagarajan, 1974, Sugiyono, 2016b).



Gambar 5. Langkah-Langkah Penelitian Pengembangan 4-D

Model 4-D yang digunakan pada penelitian ini dikembangkan lebih lanjut seperti pada gambar berikut



Gambar 6. Desain Penelitian 4-D

Berdasarkan penggunaan model pengembangan 4-D, maka tahapan pengembangan asesmen kognitif aspek HOTS dalam penelitian ini sebagai berikut.

Tabel 6. Tahap Pengembangan Model 4-D

Tahapan 4-D		Deskripsi
Tahap Pendefinisian	Analisis Kebutuhan	Melakukan analisis kebutuhan dengan menyebarkan kuesioner terhadap 41 guru peminatan kejuruan.
Tahap Perancangan	Desain Produk Awal	<ol style="list-style-type: none"> 1. Menganalisis KD yang akan dikembangkan HOTS. 2. Menyusun kisi-kisi tes. 3. Memilih stimulus yang menarik dan kontekstual. 4. Menulis 25 butir pertanyaan sesuai kisi-kisi soal. 5. Membuat kunci jawaban.
Tahap Pengembangan	Validitas Konten	Melakukan validasi oleh 5 ahli materi terkait kesesuaian butir soal dengan kisi-kisi dan hasilnya 21 butir soal dinyatakan valid.
	Revisi Produk Awal	Melakukan revisi berdasarkan oleh ahli materi, sehingga menghasilkan instrumen tes yang memenuhi kriteria valid dan dapat digunakan dalam tahap uji coba.
	Validitas Konstruk	Dilakukan uji coba terbatas 21 butir soal pada 33 peserta didik untuk menganalisis butir soal (tingkat kesukaran tes, daya beda, dan indeks pengecoh). Hasilnya 18 butir soal dinyatakan valid.
	Revisi Uji Coba Terbatas	Merevisi butir soal terkait tingkat kesukaran tes, daya beda, dan indeks pengecoh.
Tahap Penyebaran	Implementasi	Mengimplementasikan produk akhir yang telah direvisi pada tahap sebelumnya pada 87 peserta didik.

B. Populasi, Sampel, dan Tempat Penelitian

Populasi pada penelitian ini adalah seluruh peserta didik pada kompetensi keahlian Konstruksi Badan Pesawat Udara (KBPU) yang berjumlah 87 orang dan guru-guru yang mengampu mata pelajaran muatan peminatan kejuruan (C1-C3) yang berjumlah 41 orang. Teknik pengambilan sampel dalam penelitian ini menggunakan *total sampling* yaitu teknik pengambilan sampel dimana jumlah sampel sama dengan populasi, mengingat jumlah populasi yang kurang dari 100 orang (Sugiyono, 2007). Adapun tempat penelitian ini berada di salah satu SMKN di kota Bandung.

C. Instrumen dan Teknik Pengumpulan Data Penelitian

Instrumen yang digunakan dalam penelitian terdiri atas tes dan non-tes.

Tabel 7. Instrumen dan Teknik Pengumpulan Data Penelitian

No.	Jenis Data	Instrumen Pengumpulan Data	Teknik Pengumpulan Data	Sumber Data
1.	Validitas konten butir tes kognitif	- Lembar validitas isi - Lembar kisi-kisi - Lembar tes kognitif yang terdiri dari 25 butir soal PG Dua Tingkat	<i>Cheklis</i> kesesuaian antara butir tes dengan indikator	5 orang ahli konten materi
2.	Validitas konstruk dan reliabilitas butir tes	- Lembar tes kognitif yang terdiri dari 21 butir soal PG Dua Tingkat - Lembar jawaban	Melaksanakan tes tertulis secara individu	33 peserta didik
3.	Tingkat HOTS Marzano	- Lembar tes kognitif yang terdiri dari 18 butir soal PG Dua Tingkat - Lembar jawaban	Melaksanakan tes tertulis secara individu	87 peserta didik
4.	Validitas konten kuesioner	Kuesioner yang terdiri dari 16 butir pertanyaan	Melakukan uji keterbacaan	11 guru SMK
5.	Analisis kebutuhan melalui kuesioner	Kuesioner yang terdiri dari 14 butir pertanyaan	Menyebarkan kuesioner terbuka	41 guru mata pelajaran muatan peminatan kejuruan

D. Analisis Data

Teknik analisis data dilakukan terhadap pengembangan asesmen kognitif aspek HOTS, tingkat HOTS peserta didik dan kuesioner persepsi guru terkait HOTS. Beberapa teknik analisis data yang digunakan dalam penelitian ini antara lain uji validitas konten, uji validitas konstruk, uji reliabilitas, dan uji *mean*.

1. Uji Validitas Konten

Validitas isi berfungsi untuk mengecek kecocokan di antara butir-butir soal dengan indikator, materi, atau tujuan pembelajaran yang telah ditetapkan oleh mayoritas ahli (Susetyo, 2015). Dalam penelitian ini, format yang digunakan

untuk validasi ahli yaitu format dikotomi dengan kriteria penilaian sebagai berikut.

Tabel 8. Kriteria Penilaian Instrumen

Kriteria	Bobot
Sesuai	1
Tidak Sesuai	0

Setelah data validitas berdasarkan penilaian para ahli terkumpul, maka data tersebut dihitung dengan beberapa cara, yaitu:

a. Persentase butir yang cocok dengan indikator/tujuan

Dalam konteks penelitian ini, cara perhitungan ini digunakan untuk melihat aspek kecocokan (kesesuaian) antara butir soal dengan kisi-kisi tes kognitif (mencakup KD, materi, indikator soal dan level kognitif). Butir tes dinyatakan valid jika kecocokannya dengan indikator mencapai >50%.

Rumus yang digunakan sebagai berikut.

$$\text{Persentase} = \frac{f}{\sum f} \times 100\%$$

(Sumber: Susetyo, 2015).

Keterangan:

f : Frekuensi sesuai menurut penilai
 $\sum f$: Jumlah penilai

b. Perhitungan rasio validitas isi dari C H Lawshe

Dalam konteks penelitian ini, rumus CVR digunakan untuk melihat kesesuaian antara butir soal dengan aspek materi pada instrumen telaah soal HOTS bentuk PG yang diadaptasi dari Kemendikbud (2017). Rumus yang digunakan sebagai berikut.

$$CVR = \frac{Ne - N/2}{N/2}$$

(Sumber: Gilbert & Prion, 2016)

Keterangan:

CVR : Rasio validitas konten
 Ne : Jumlah validator yang menyatakan setuju
 N : Total validator

Hasil perhitungan CVR akan dibandingkan dengan nilai kritis CVR Lawshe pada tabel berikut dengan ketentuan:

Bila $CVR_{hitung} > CVR_{tabel}$, maka butir soal valid

Bila $CVR_{hitung} < CVR_{tabel}$, maka butir soal tidak valid

Tabel 9. Nilai Kritis CVR Lawshe

Jumlah Ahli	Nilai CVR Minimum
5	0,573
6	0,523
7	0,485
8	0,453
9	0,427
10	0,405
11	0,387

(Sumber: Wilson, Pan, & Schumsky, 2012)

2. Uji Validitas Konstruk

Baik buruknya tes tergantung pada butir-butir soal yang ada di dalamnya. Oleh sebab itu, untuk mendapatkan tes yang baik perlu dipilih butir-butir yang baik, sedangkan butir yang buruk harus dibuang dan yang kurang baik perlu direvisi. Untuk mengetahui kualitas tiap butir soal perlu dianalisis satu persatu. Analisis meliputi perhitungan daya beda, tingkat kesukaran, homogenitas tes serta analisis distraktor/pengecoh pada tes.

a. Daya Beda (DB)

Daya beda (*item discrimination*) adalah indeks yang digunakan untuk membedakan peserta tes berkemampuan tinggi (kelompok atas) dan berkemampuan rendah (kelompok bawah) (Toksöz & Ertunç, 2017). Tahapan pertama dalam menghitung DB adalah menentukan kelompok atas dan kelompok bawah. Pada umumnya, para ahli tes membagi kelompok ini menjadi 27% atau 33% kelompok atas dan 27% atau 33% kelompok bawah. Namun yang paling stabil dan sensitive serta paling banyak digunakan adalah dengan menentukan 27% kelompok atas dan 27% kelompok bawah, sisa lembar jawaban sebanyak 46% disisihkan (Surapranata, 2006) (Mukherjee & Lahiri, 2015). Rumus dalam menghitung DB masing-masing *item*:

$$DP = \frac{Ba}{na} - \frac{Bb}{nb}$$

(Sumber: Toksöz & Ertunç, 2017, Surapranata, 2006)

Keterangan :

- DP : Indeks daya pembeda
 Ba : Jumlah peserta didik yang menjawab soal dengan benar pada kelompok atas
 Bb : Jumlah peserta didik yang menjawab soal dengan benar pada kelompok bawah
 na : Jumlah peserta didik kelompok atas
 nb : Jumlah peserta didik kelompok bawah

Proses interpretasi daya pembeda seperti yang disajikan pada tabel 9.

Tabel 10. Kriteria Daya Beda

Interval	Kualitas Item
< 0,20	Buruk
0,20 – 0,40	Cukup
> 0,40	Baik

(Sumber: Basuki & Hariyanto, 2016)

b. Tingkat Kesukaran (TK)

Tingkat kesukaran (*Difficulty Index*) suatu butir soal ditentukan kedalaman soal, kompleksitas, atau hal-hal lain yang berkaitan dengan kemampuan yang diukur oleh soal (Surapranata, 2006). Walaupun kita melakukan kajian lebih dalam, akan terasa sulit dalam menentukan sebuah soal lebih sukar dibandingkan dengan soal yang lain. Para ahli tes kemudian menentukan formula dalam menentukan tingkat kesukaran yakni mengacu pada banyaknya peserta tes yang dapat menjawab benar pada suatu soal. Jawaban yang benar diberi skor 1, sedangkan jawaban yang salah diberi skor 0. Rumus yang digunakan untuk menghitung tingkat kesukaran yaitu:

$$P = \frac{B}{JS}$$

(Sumber: Toksöz & Ertunç, 2017, Surapranata, 2006)

Keterangan:

- P : tingkat kesukaran perangkat tes
 B : Banyaknya peserta tes yang menjawab soal dengan benar
 JS : Jumlah seluruh peserta tes

Proses interpretasi tingkat kesukaran seperti yang disajikan pada tabel 10.

Tabel 11. Kriteria Tingkat Kesukaran

Interval	Kualitas Item
< 0,30	Sukar
0,30 - 0,70	Sedang
> 0,70	Mudah

(Sumber: Rao, L, Sajitha, & Permi, 2016, Kolte, 2015)

c. Indeks Pengecoh (*Distractor Efficiency*)

Pengecoh non fungsional (*Non-Functional Distractors*) adalah opsi yang jarang dipilih (<5%) oleh peserta tes dan pengecoh fungsional (efektif) adalah pilihan yang dipilih oleh 5% atau lebih peserta tes (Mukherjee & Lahiri, 2015). Dengan demikian, pengecoh non fungsional seharusnya direvisi, dihapus atau diganti dengan opsi yang lebih masuk akal. Efisiensi pengecoh ditentukan untuk setiap item berdasarkan jumlah NFD di dalamnya dan berkisar dari 0 hingga 100%.

$$px_i = \frac{fx_i}{M} \times 100\%$$

Keterangan:

px_i : proporsi masing-masing pilihan jawaban suatu butir tes

fx_i : frekuensi masing-masing pilihan jawaban suatu butir tes

M : jumlah responden

Proses interpretasi indeks pengecoh seperti yang disajikan pada tabel berikut.

Tabel 12. Kriteria Indeks Pengecoh

Interval	Tingkat Hubungan
>0,05	Distraktor berfungsi/efektif
<0,05 atau 5%	Distraktor tidak berfungsi

(Sumber: Rao et al., 2016)

2. Uji Reliabilitas Instrumen

Instrumen tidak hanya harus valid, namun juga harus reliabel. Untuk menguji reliabilitas instrumen dapat dilakukan dengan banyak metode, salah satunya yaitu Metode Kuder Richardson (KR). Metode ini menghitung secara langsung seluruh butir tes, tanpa membagi menjadi dua bagian pada perangkat ukur. Apabila semua butir tes setara, maka koefisien reliabilitas KR_{20} adalah

$$\rho_{KR20} = \frac{k}{k-1} \left[1 - \frac{\sum pq}{\sigma_x^2} \right]$$

Fendy Thomas, 2019

PENGEMBANGAN ASESMEN KOGNITIF ASPEK HIGHER ORDER THINKING SKILL PADA PEMBELAJARAN VOKASI

Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu

Untuk mencari nilai varian skor tes menggunakan rumus sebagai berikut.

$$\sigma_x^2 = \frac{N\sum X^2 - (\sum X)^2}{N^2}$$

(Sumber: Susetyo, 2015)

Keterangan:

- p : Proporsi jawaban benar
 q : Proporsi jawaban salah
 k : Jumlah butir tes
 $\sum pq$: Jumlah perkalian jawaban benar dan salah
 ρ_{KR20} : Koefisien reliabilitas
 σ_x^2 : Varian skor tes
 N : Jumlah responden

Interpretasi nilai reliabilitas suatu instrumen dapat dikategorikan dalam tabel 12.

Tabel 13. Kriteria Reliabilitas Butir Soal

Koefisien reliabilitas	Interpretasi
$0.90 < r_{11} \leq 1.00$	sangat tinggi
$0.70 < r_{11} \leq 0.90$	tinggi
$0.40 < r_{11} \leq 0.70$	sedang
$0.20 < r_{11} \leq 0.40$	rendah
$r_{11} \leq 0.20$	sangat rendah

(Sumber: Arikunto, 2014)

4. Uji Mean

a. Tingkat HOTS Marzano

Tingkat HOTS Marzano pada peserta didik diketahui dengan cara melakukan uji statistika *mean* dengan formula:

$$Mean = \frac{\sum X}{n}$$

Proses interpretasi skor *mean* seperti yang disajikan pada tabel berikut.

Tabel 14. Kriteria Mean

Skor Mean	Tingkat
< 0,54	Rendah
0,54 – 1,45	Sedang
> 1,45	Tinggi

Keterangan:

Mean : Rata-rata

$\sum X$: Jumlah Data

n : Banyak data

b. Persepsi Guru terhadap HOTS

Persepsi guru terhadap HOTS juga diketahui dengan cara:

$$PPS (\%) = \frac{NS}{N} \times 100\%$$

Ket:

PPS (%) : Persentase Persetujuan guru terhadap suatu pertanyaan

Ns : Jumlah guru yang menyatakan setuju

N : Jumlah seluruh peserta didik

Proses interpretasi persentase disajikan pada tabel berikut.

Tabel 15. Interpretasi Persepsi Guru

Persentase	Kategori
0%	Tak seorangpun
1% - 25%	Sebagian kecil
26% - 49%	Hampir separuhnya
50%	Separuhnya
51% - 75%	Sebagian besar
76% - 99%	Hampir seluruhnya
100%	Seluruhnya

(Sumber: Arikunto, 2014)

Kuesioner divalidasi oleh 11 guru. Validitas isi dilakukan untuk melihat isi produk awal yang bertujuan untuk mendapatkan masukan, saran perbaikan, dan sekaligus penilaian terhadap kuesioner. Berikut data hasil validitas ahli.

Tabel 16. Data Butir Kuesioner Hasil Validitas Konten

Nomor Butir Pertanyaan	Jumlah Butir	Nilai CVR	Kriteria
1, 2, 3, 4, 5, 6, 7, 8, 10, 11, 12, 13, 14, dan 15	14	>0,387	Valid
9 dan 16	2	<0,387	Tidak Valid

Ket: Kisi-kisi kuesioner lengkap lihat di lampiran 2

Pada tabel di atas, dijelaskan pengolahan data dari hasil *judgment*. Butir nomor 9 dan 16 merupakan butir pertanyaan yang diolah berdasarkan pendapat *judgment* terkait keterbacaan, ketepatan, kemudahan maupun kesulitan pemahaman terhadap butir pertanyaan tersebut. Dari total 11 guru, tujuh guru

merespon “Sesuai” dan empat guru merespon “Tidak Sesuai” terhadap butir pertanyaan nomor 9 dan 16. Kemudian dilakukan perhitungan CVR, dimana hasil CVR_{hitung} dibandingkan dengan CVR_{tabel} , sehingga bila mengacu pada Nilai Kritis Lawshe (pada tabel 9), maka *item* tersebut termasuk kategori tidak valid dan akan dihilangkan. Kuesioner yang dinyatakan valid berjumlah 14 butir.