

PENGEMBANGAN INSTRUMEN ANGKET

Buku berjudul "PENGEMBANGAN INSTRUMEN ANGKET" merupakan buku monograf yang mengkaji tentang bagaimana cara menyusun dan mengembangkan instrumen angket yang siap digunakan dalam penelitian menggunakan desain kuantitatif. Buku ini disusun dari berbagai referensi dan hasil penelitian. Selain itu, berbagai artikel hasil penelitian berisi isu lingkungan terkait pengembangan instrumen angket yang bersumber dari artikel dalam lingkup nasional maupun internasional juga diangkat guna melengkapi dan meng-update isi buku.

Penerbit K-Media
Bantul, Yogyakarta
kmediacorp
kmedia.cv@gmail.com
www.kmedia.co.id



PENGEMBANGAN INSTRUMEN ANGKET

Dr. Rohmad, M.Pd. & Dr. Siti Sarah, M.Pd.



PENGEMBANGAN INSTRUMEN ANGKET

Dr. Rohmad, M.Pd.
Dr. Siti Sarah, M.Pd.

PENGEMBANGAN INSTRUMEN ANGKET

Dr. Rohmad, M.Pd.

Dr. Siti Sarah, M.Pd.



Penerbit K-Media
Yogyakarta, 2021

PENGEMBANGAN INSTRUMEN ANGKET

vi + 133 hlm.; 15,5 x 23 cm

ISBN: 978-623-316-317-0

Penulis : Rohmad & Siti Sarah

Editor : Ali Murtadho

Tata Letak : Nur Huda A.

Desain Sampul : Nur Huda A.

Cetakan 1 : September 2021

Copyright © 2021 by Penerbit K-Media
All rights reserved

Hak Cipta dilindungi Undang-Undang No 19 Tahun 2002.

Dilarang memperbanyak atau memindahkan sebagian atau seluruh isi buku ini dalam bentuk apapun, baik secara elektris maupun mekanis, termasuk memfotocopy, merekam atau dengan sistem penyimpanan lainnya, tanpa izin tertulis dari Penulis dan Penerbit.

Isi di luar tanggung jawab percetakan

Penerbit K-Media
Anggota IKAPI No.106/DIY/2018
Banguntapan, Bantul, Yogyakarta.
e-mail: kmedia.cv@gmail.com

KATA PENGANTAR

Puji syukur kami panjatkan kepada Tuhan Yang Maha Esa atas anugerah-Nya sehingga penulisan buku monograf “Pengembangan Instrumen Angket” dapat diselesaikan dengan baik. Buku ini tercipta sebagai keinginan penulis untuk menyediakan buku yang mengkaji materi Metodologi Penelitian Kuantitatif Pendidikan dalam bentuk yang ringan dan simpel sesuai dengan karakteristik mahasiswa di tingkat strata 1 khususnya di Fakultas Tarbiyah dan Ilmu Keguruan, IAIN Purwokerto. Buku monograf juga dilengkapi dengan contoh kasus berupa penerapan isu lingkungan dalam bentuk hasil penelitian. Melalui buku ini, penulis berharap mahasiswa akan lebih mudah memahami materi sekaligus mempraktekkan penyusunan instrumen angket sebagai bekal dalam menyelesaikan tugas akhir. Karya ini juga tercipta sebagai bentuk aktualisasi pada kegiatan Latihan Dasar CPNS 2021 yang diselenggarakan oleh Balai Diklat Keagamaan Semarang.

Penyelesaian buku ini tidak terlepas dari bantuan berbagai pihak. Untuk itu, ucapan terima kasih penulis sampaikan kepada berbagai pihak yang telah mendukung penyelesaian buku ini melalui berbagai peran.

1. Bapak Drs. H. Anshori selaku Kepala Balai Pendidikan dan Pelatihan Keagamaan Semarang
2. Bapak Dr. Moh. Roqib, M. Ag. selaku Rektor IAIN Purwokerto.
3. Ibu Siti Kusriyah, M.Pd. selaku *coach* yang telah banyak membimbing sehingga tercipta buku ini.
4. Bapak Tri Wahyono selaku mentor yang telah memberi motivasi dan bimbingan.
5. Panitia Latsar CPNS Angkatan III Balai Diklat Keagamaan Semarang yang telah merancang dan memfasilitasi kegiatan Latihan Dasar CPNS Golongan III.
6. Para Widyaiswara Balai Diklat Keagamaan Semarang yang telah memberikan ilmu dan teladan kepada peserta Latsar.

7. Rekan-rekan peserta Latsar CPNS 2021 khususnya Golongan III yang telah memberi dukungan.

Kami selaku penulis menyadari bahwa buku monograf ini belum sempurna. Oleh karena itu, kritik dan saran dari pembaca sangat diharapkan untuk menyempurnakan kualitas buku ini. Penulis sangat berharap semoga buku ini memiliki manfaat terhadap eksistensi dan perkembangan ilmu pengetahuan khususnya dalam pembelajaran metodologi penelitian kuantitatif pendidikan.

Purwokerto, September 2021
Penulis,

Tim

DAFTAR ISI

KATA PENGANTAR	iii
DAFTAR ISI.....	v
BAB I PARADIGMA PENELITIAN.....	1
A. Pengertian Penelitian.....	1
B. Paradigma Penelitian Kuantitatif dan Kualitatif	6
BAB II INSTRUMEN PENELITIAN.....	11
A. Pengertian	11
B. Penilaian Afektif	12
C. Angket.....	16
1. Angket terbuka.....	17
2. Angket tertutup	17
D. Skala Penilaian.....	21
1. Skala Likert.....	23
2. Skala Guttman.....	26
3. Skala diferensial semantik (<i>Semantic</i> <i>Differensial Scale</i>)	27
4. Rating skala (skala penilaian)	28
5. Skala Thurstone	29
BAB III PENGEMBANGAN INSTRUMEN	31
A. Langkah Pengembangan Instrumen	31
B. Contoh Pengembangan Instrumen.....	42
BAB IV VALIDITAS INSTRUMEN.....	76
A. Pengertian	76
B. Validitas Isi dan Cara Pembuktiannya.....	77
C. Validitas Konstruk dan Cara Pembuktiannya.....	89

BAB V ANALISIS HASIL UJICOBA INSTRUMEN	
ANGKET	103
A. Prosedur Ujicoba Instrumen Angket	103
B. Analisis Hasil Ujicoba Instrumen Angket	103
1. Validitas empiris	106
2. Reliabilitas Angket.....	112
DAFTAR PUSTAKA	123
BIODATA PENULIS	130

BAB I

PARADIGMA PENELITIAN

A. Pengertian Penelitian

Research is a process of steps used to collect and analyze information to increase our understanding of a topic or issue (Creswell: 2012). Artinya, penelitian merupakan sebuah proses yang terdiri atas beberapa langkah, yaitu mengumpulkan dan menganalisis informasi untuk meningkatkan pemahaman kita tentang suatu topik atau masalah. Definisi lain disampaikan oleh Kerlinger (2002), yang mendefinisikan penelitian sebagai penelitian yang sistematis, terkontrol, empiris, dan penyelidikan kritis dari proporsi-proporsi hipotesis tentang hubungan yang diperkirakan antara gejala alam. Berdasarkan kedua pendapat di atas, dapat disimpulkan bahwa penelitian merupakan suatu proses penyelidikan terhadap suatu masalah secara sistematis, kritis, ilmiah, dan teliti untuk mencari jawaban atau memecahkan suatu masalah.

Creswell (2012) menyatakan memuat 3 langkah pokok penelitian, yaitu (1) mengajukan pertanyaan, (2) mengumpulkan data untuk menjawab pertanyaan tersebut, dan (3) menyajikan jawaban atas pertanyaan. Artinya, penelitian diawali dari sebuah pertanyaan terkait penyelesaian masalah. Untuk menyelesaikan masalah tersebut diperlukan data yang relevan. Hasil analisis data digunakan untuk menjawab pertanyaan penelitian.

Sudaryono, Margono, dan Rahayu (2013) membagi tahapan pelaksanaan penelitian dalam 8 langkah yaitu (1) mengidentifikasi masalah, (2) merumuskan dan membatasi masalah, (3) melakukan studi kepustakaan, (4) merumuskan hipotesis atau pertanyaan penelitian, (5) menentukan desain dan metode penelitian, (6) menyusun instrumen dan mengumpulkan data, (7) menganalisis data dan menyajikan hasil, dan (8) menginterpretasikan temuan, membuat kesimpulan, dan saran. Berikut akan dijelaskan lebih detail.

Sebuah penelitian dimulai dengan mengidentifikasi isu-isu dan masalah yang penting, aktual, mendesak, serta memiliki banyak manfaat. Perumusan masalah bertujuan untuk memfokuskan tujuan penelitian yang akan dicapai berdasarkan variabel penelitian yang dituju.

Studi kepustakaan diperlukan guna menentukan teori yang melandasi variabel penelitian sekaligus metodologi penelitian yang akan dipakai dalam penelitian. Kajian terkait hasil penelitian relevan akan sangat bermanfaat sebagai salah satu referensi guna memperajam kajian yang akan dibahas.

Perumusan hipotesis bertujuan untuk menduga hasil penelitian berdasarkan teori dan referensi penelitian. Hipotesis banyak digunakan dalam penelitian kuantitatif, sebaliknya penelitian kualitatif tidak perlu menggunakan hipotesis. Sebagai gantinya, penelitian kualitatif cukup menggunakan pertanyaan penelitian. Langkah lanjutan dari penentuan hipotesis adalah menentukan desain dan metode penelitian, meliputi langkah penelitian, metode penelitian, dan sumber data.

Menyusun instrumen dan mengumpulkan data menjadi hal penting dalam suatu penelitian bahkan menjadi hal terpenting dalam penentuan baik tidaknya hasil penelitian. Penelitian yang baik akan terpenuhi jika instrumen yang digunakan sudah memenuhi standar instrumen yang baik. Standar itu meliputi validitas dan reliabilitas instrumen. Validitas merujuk pada ketepatan instrument dalam mengukur hal yang harus diukur. Adapun reliabilitas merujuk pada keajegan instrumen dalam menilai responden. Kualitas instrumen yang baik akan sangat berpengaruh terhadap kualitas penelitian.

Analisis data menjelaskan teknik dan langkah peneliti dalam mengolah dan menganalisis data. Rumusan masalah menjadi salah satu pertimbangan dalam menentukan teknik analisis data penelitian. Hal ini dikarenakan analisis data berkaitan erat dengan cara seorang peneliti dalam menjawab rumusan masalah penelitian. Hasil analisis data penelitian belumlah cukup tanpa pemberian makna melalui interpretasi data. Interpretasi data mempermudah orang lain dalam memahami hasil penelitian.

Poin penting dari analisis data beserta interpretasinya yaitu pengambilan kesimpulan berisi jawaban terhadap rumusan penelitian yang telah ditetapkan. Di bagian akhir penelitian, peneliti memberikan saran yang dapat berisi rekomendasi terhadap pihak-pihak terkait dalam melaksanakan dan memanfaatkan penelitian.

Khusus dalam bidang pendidikan, Creswell (2012) menyampaikan 6 langkah pelaksanaan penelitian pendidikan, yaitu: (1) identifikasi masalah penelitian, (2) melakukan kajian pustaka, (3) menentukan tujuan penelitian, (4) mengumpulkan data, (5) menganalisis dan menafsirkan data, (6) melaporkan dan mengevaluasi penelitian. Identifikasi masalah merupakan suatu kegiatan berupa penentuan isu atau masalah yang akan diselesaikan melalui sebuah penelitian. Hal ini dilakukan melalui beberapa langkah, seperti menentukan masalah dan menyarankan pentingnya penelitian untuk khalayak yang akan membaca hasil penelitian. Melalui langkah ini, peneliti akan lebih fokus dalam melakukan penelitian.

Kajian pustaka merupakan langkah penelitian yang bertujuan mempelajari masalah penelitian secara lebih mendalam sekaligus mencari dasar pengetahuan yang akan digunakan dalam menyelesaikan masalah. Ini sangat penting salah satunya agar penelitian yang dilakukan tidak hanya sebuah replika penelitian sebelumnya.

Kajian literatur memuat serangkaian kegiatan yaitu: (1) menemukan ringkasan, buku, jurnal, dan publikasi terindeks pada suatu topik; (2) selektif memilih literatur yang akan disertakan dalam ulasan; dan (3) meringkas literatur dalam laporan tertulis. Hal ini penting diperlukan guna menentukan teori yang menjadi dasar pelaksanaan penelitian.

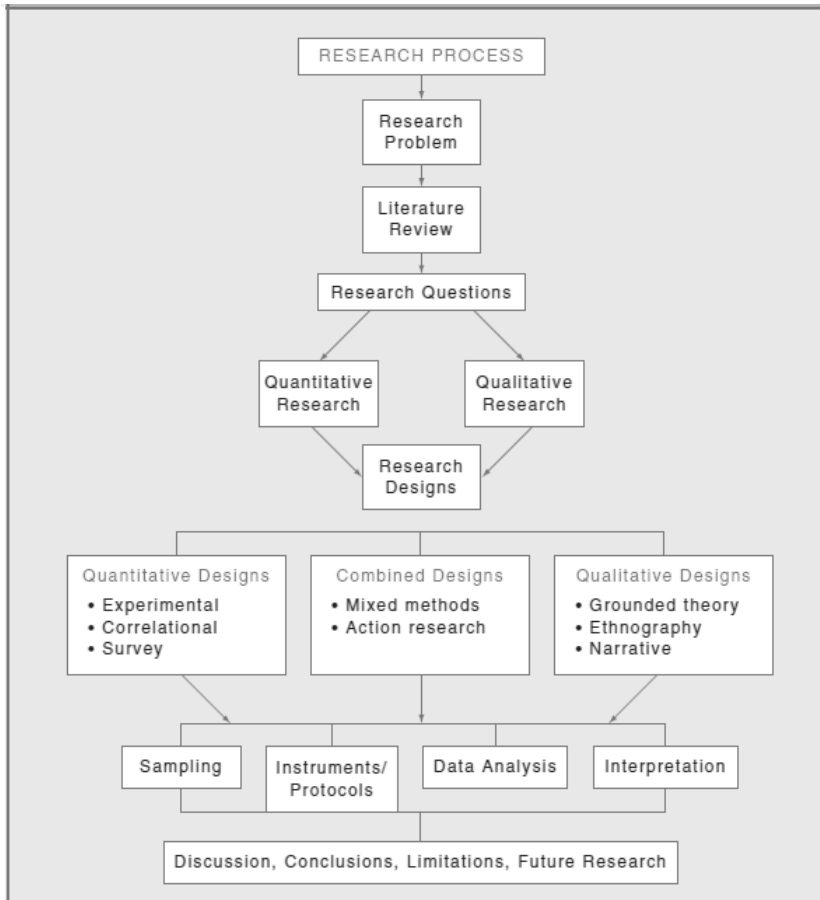
Penentuan tujuan penelitian menjadi langkah peneliti dalam menentukan tujuan penelitian secara spesifik guna memfokuskan penelitian. Melalui tujuan penelitian, peneliti akan lebih mudah memfokuskan langkah penelitian.

Mengumpulkan data berarti mengidentifikasi dan memilih responden untuk penelitian, mendapatkan izin responden untuk mempelajarinya, dan mengumpulkan informasi dengan mengajukan pertanyaan atau mengamati perilaku responden. Perhatian utama proses ini adalah kebutuhan untuk memperoleh data yang akurat. Data yang dikumpulkan dapat berupa angka (seperti skor tes, frekuensi perilaku) atau kata-kata (seperti tanggapan, pendapat, kutipan). Jenis data akan sangat bergantung pada instrument yang akan dipilih.

Tindak lanjut dari pengumpulan data adalah melakukan analisis data untuk menentukan tanggapan responden sekaligus menafsirkannya. Menganalisis dan menafsirkan data melibatkan penarikan kesimpulan penelitian, merepresentasikannya dalam tabel dan gambar untuk meringkasnya, dan menjelaskan kesimpulan dengan kata-kata untuk memberikan jawaban atas pertanyaan penelitian.

Tahap akhir pelaksanaan penelitian pendidikan adalah pelaporan dan evaluasi penelitian. Dalam menuliskan laporan, peneliti harus menghormati dan menghindari bahasa yang mendiskriminasi berdasarkan jenis kelamin, ras, atau kelompok etnis. Evaluasi penelitian dapat dilakukan dengan melakukan desiminasi ke berbagai pihak atau audiens yang kompeten di bidang yang dikaji peneliti.

Lebih lanjut, Creswell (2012) menjelaskan langkah penelitian pendidikan yang didalamnya mengakomodir desain penelitian seperti terlihat pada gambar berikut.



Gambar 1.1 Proses penelitian
(Sumber: Creswell, 2012)

Gambar 1.1 menggambarkan bahwa penelitian dimulai dari sebuah masalah. Masalah yang ada perlu dilakukan tinjauan dengan mendasarkan pada berbagai teori dan referensi yang sesuai. Langkah ini dilakukan agar peneliti lebih fokus dalam memilih pertanyaan penelitian yang menjadi fokus penelitian. Pertanyaan penelitian akan mengarah pada jenis penelitian, apakah kualitatif atau kuantitatif. Langkah ini akan berimbas pada penentuan desain penelitian. Berdasarkan desain penelitian yang digunakan, penelitian terbagi

menjadi 3 jenis, yaitu desain penelitian kuantitatif, penelitian kualitatif, dan penelitian campuran kuantitatif dan kualitatif (*mix methods*). Penentuan desain penelitian akan merujuk pada pemilihan teknik sampling, jenis instrumen, analisis data, dan interpretasi data. Penentuan berbagai hal secara sinkron akan sangat menentukan hasil penelitian yang dikaji lebih jauh dalam bentuk diskusi, kesimpulan, keterbatasan, dan riset lanjutan. Untuk lebih memahami jenis penelitian kuantitatif dan kualitatif, berikut akan disajikan bahasan terkait paradigma penelitian kuantitatif dan kualitatif.

B. Paradigma Penelitian Kuantitatif dan Kualitatif

Secara garis besar paradigma penelitian terbagi menjadi 2 macam, yaitu penelitian kuantitatif dan penelitian kualitatif. Karakteristik penelitian kuantitatif menurut Purwanto (2010) dapat dilihat dari 7 point yaitu (1) dipengaruhi metode penelitian alam, (2) bersifat behavioristik-mekanistik-empirik, (3) memberikan perhatian pada hasil, (4) tujuan penelitian adalah untuk mendapatkan aturan, hukum, dan prinsip yang bersifat umum, (5) konversi kualitas menjadi kuantitas, (6) konfirmasi teori, dan (7) menjunjung teori objektivitas. Masing-masing akan dibahas satu persatu. Penjelasan masing-masing point terdapat pada bahasan berikut.

Pendekatan kuantitatif dipengaruhi metode penelitian alam terlihat dari pengaruh ilmu dalam metode observasi yang diambil dari cara yang digunakan dalam ilmu astronomi, asumsi normal dalam pengukuran psikologis, dan sebagainya. Selain itu, penelitian kuantitatif bersifat behavioristik-mekanistik-empirik (Dahar, 1988). Teori belajar behavioristik mendefinisikan belajar sebagai perubahan perilaku yang dapat diamati melalui kaitan antara stimulus dan respon (S-R) menurut prinsip yang mekanistik. Para behavioris meyakini bahwa hasil belajar akan lebih baik dikuasai kalau dilakukan secara berulang-ulang. Jadi, belajar tidak memerlukan pemahaman karena terbentuk akibat ikatan S-R yang diperkuat dengan pengulangan.

Para peneliti kuantitatif memberikan perhatian pada hasil belajar. Hal ini terlihat dari cara yang dilakukan para peneliti kuantitatif dalam memahami manusia dengan melihat apa yang telah dihasilkannya,

dalam bentuk hubungan atau kausalitas antargelaja pada diri manusia. Penelitian kuantitatif tidak memfokuskan pada keunikan tapi menyusun aturan, hukum, dan prinsip umum tentang dunia yang ditemukan dalam data empiris menggunakan sampe yang luas. Artinya, tujuan penelitian kuantitatif lebih banyak mengarah ke pembuatan hukum dari generalisasi.

Penelitian kuantitatif mendasarkan penelitian pada data berbentuk angka. Artinya, ada konversi antara kualitas menuju kuantitas dalam menilai variabel penelitian. Perkembangan instrumen penelitian yang searah perkembangan ilmu menuntut peneliti kuantitatif untuk selalu meng-upgrade keterampilan terkait penyusunan instrumen penelitian.

Pada penelitian kuantitatif, teori menjadi panduan untuk menguji kebenaran empiris. Teori bukan hanya sebuah kebenaran, tetapi menjadi isyarat adanya kebenaran yang harus diverifikasi.

Empiris sebagai sifat dari penelitian kuantitatif berimbas pada sikap peneliti kuantitatif yang menjunjung tinggi objektivitas, bahkan menganggapnya sebagai salah satu persyaratan dasar pengetahuan yang benar.

Desain penelitian kuantitatif meliputi survey, penelitian korelasi, desain penelitian kausal komparatif, dan eksperimen. Survei adalah prosedur dalam penelitian kuantitatif di mana peneliti mengelola survei kepada sekelompok kecil orang (disebut sampel) untuk mengidentifikasi tren dalam sikap, pendapat, perilaku, atau karakteristik sekelompok besar orang. Saat pandemi Covid-19 melanda, survey menjadi salah satu desain penelitian. Demikian karakteristik penelitian kuantitatif.

Penelitian kualitatif adalah kegiatan pengumpulan, analisis, dan interpretasi data naratif dan visual yang komprehensif (yaitu, nonnumeric atau bukan angka) untuk mendapatkan wawasan tentang fenomena tertentu yang menarik. Artinya, data bukan angka menjadi salah satu ciri khusus penelitian kualitatif. Selain itu, peneliti kualitatif tidak serta merta menerima pandangan dunia yang stabil, koheren, dan seragam. Semua makna terletak dalam perspektif atau konteks tertentu, dan karena orang dan kelompok yang berbeda sering kali memiliki perspektif dan konteks yang berbeda, dunia memiliki banyak

makna yang berbeda. Artinya, tidak ada yang lebih valid atau benar daripada yang lain. Untuk lebih memahami perbedaan antara penelitian kualitatif dan kuantitatif, perhatikan Tabel 1. Adapun perbedaan kedua desain jika dikaitkan dengan tahapan penelitian seperti dapat dilihat pada Tabel 2.

Tabel 1.1 Perbedaan penelitian kuantitatif dan kualitatif

(Sumber: Gay, Milles, & Airasian, 2012)

Perbedaan	Penelitian Kuantitatif	Penelitian Kualitatif
Jenis data	Data numeric (angka)	Narasi nonnumerik dan data visual
Masalah dalam penelitian	Hipotesis dan prosedur penelitian dinyatakan sebelum memulai studi	Masalah dan metode penelitian berkembang saat pemahaman topik semakin dalam
Manipulasi konteks	Ya	Tidak
Jumlah sampel	Lebih banyak	Lebih kecil
Prosedur penelitian	Bergantung pada prosedur statistik	Mengandalkan pengkategorian dan pengorganisasian data ke dalam pola untuk menghasilkan sintesis deskriptif dan naratif
Interaksi peserta	Sedikit	Banyak
Keyakinan yang mendasari	Kita hidup di dunia yang stabil dan dapat diprediksi yang dapat kita ukur, pahami, dan umumkan.	Makna terletak dalam perspektif atau konteks tertentu yang berbeda bagi orang dan kelompok; oleh karena itu, dunia memiliki banyak arti.

Tabel 1.2 Perbedaan karakteristik penelitian kuantitatif dan kualitatif berdasarkan tahapan proses penelitian

(Sumber: Gay, Milles, & Airasian, 2012)

Karakteristik penelitian kuantitatif	Proses penelitian	Karakteristik penelitian kualitatif
Berorientasi pada deskripsi dan penjelasan	Identifikasi masalah penelitian	Berorientasi pada eksplorasi dan pemahaman
<ul style="list-style-type: none"> ✓ Peran utama ✓ Pembeneran untuk masalah penelitian dan spesifikasi untuk kebutuhan penelitian 	Melakukan kajian pustaka,	<ul style="list-style-type: none"> ✓ Peran kecil ✓ Pembeneran untuk masalah penelitian
<ul style="list-style-type: none"> ✓ Spesifik dan sempit ✓ Data yang terukur dan dapat diamati 	Menentukan sampel penelitian	<ul style="list-style-type: none"> ✓ Umum dan luas ✓ Pengalaman peserta
<ul style="list-style-type: none"> ✓ Instrumen yang telah ditentukan ✓ Data numerik ✓ Responden banyak 	Pengumpulan data	<ul style="list-style-type: none"> ✓ Protokol yang muncul ✓ Data teks atau gambar ✓ Responden sedikit
Analisis statistik <ul style="list-style-type: none"> ✓ Deskripsi tren, perbandingan kelompok, atau hubungan antar variabel ✓ Perbandingan hasil dengan prediksi dan studi sebelumnya 	Menganalisis dan menafsirkan data	Analisis teks <ul style="list-style-type: none"> ✓ Deskripsi, analisis, dan pengembangan tematik ✓ Makna yang lebih besar dari temuan
<ul style="list-style-type: none"> ✓ Standar dan tetap ✓ Objektif dan tidak bias 	Melaporkan dan mengevaluasi penelitian	<ul style="list-style-type: none"> ✓ Fleksibel dan muncul ✓ Reflektif dan bias

Berdasarkan Tabel 1.1 dan Tabel 1. 2 dapat diketahui bahwa perbedaan penelitian kualitatif dan kuantitatif sangat terlihat dari

mulai tahapan pengumpulan data. Sebagai bahan informasi, penyusunan instrumen merupakan tahapan penelitian yang memiliki peran penting dalam menentukan kualitas penelitian secara keseluruhan. Pembahasan terkait instrumen penelitian secara lebih detail dapat dilihat pada BAB II.

Contoh 1.1

Sebuah penelitian bertujuan untuk mengetahui permasalahan yang dialami guru SD dalam melakukan pembelajaran online pada masa pandemi Covid-19. Penelitian menggunakan metode survey. Hasil penelitian menunjukkan bahwa 100% guru melakukan pembelajaran daring. Media pembelajar daring yang digunakan oleh guru SD di Kabupaten Bantul, yaitu WhatsApps, WhatsApp Web, Google Classroom, Google Group, TeamLink, Microsoft Teams, Kaizala Microsoft, Zoom Meeting & Webinar, Youtube, Google Hangouts, dan lain-lain. Urutan pilihan media pembelajaran daring yang digunakan guru secara berurutan dari yang paling banyak diminati hingga yang kurang diminati, Google Class Google form, YouTube, Video pendek terkait materi, dan Zoom Cloud Meeting.

Sumber: Anugrahana, A. (2020). Hambatan, Solusi dan Harapan: Pembelajaran Daring Selama Masa Pandemi Covid-19 Oleh Guru Sekolah Dasar. *Scholaria: Jurnal Pendidikan dan Kebudayaan*, 10(3): 282-289

Penerapan Isu lingkungan 1.1

Sebuah penelitian bertujuan untuk mengetahui bagaimana metode pembelajaran, kualitas, dan tantangan dalam pembelajaran online instruktur teknik industri di Indonesia yang beralih dari pembelajaran tatap muka menjadi sepenuhnya online selama pandemi COVID-19. Penelitian menggunakan desain survey. Untuk mengetahui hasil penelitian, silahkan langsung cek di <https://doi.org/10.18870/hlrc.v11i1.1232>.

BAB II

INSTRUMEN PENELITIAN

A. Pengertian

Instrumen penelitian merupakan salah satu hal penting dalam penelitian untuk mengukur ketercapaian variabel penelitian. Hasil pengukuran yang baik akan menghasilkan data yang baik. Data hasil penelitian selanjutnya akan diolah untuk menghasilkan informasi terkait tujuan penelitian. Informasi tersebut sangat berguna untuk menyelesaikan masalah bahkan dapat menjadi salah satu sumber informasi dalam penentuan kebijakan.

Instrumen penelitian yang baik harus memenuhi uji validitas dan reliabilitas. Uji validitas terkait dengan jaminan bahwa instrumen mampu mengukur sesuatu yang hendak diukur. Adapun reliabilitas mengacu pada tingkat keajegan instrumen dalam mengukur variabel penelitian.

Penentuan validitas instrumen dimulai dengan penentuan jenis instrumen dalam suatu penelitian. Jenis instrumen terkait erat dengan variabel penelitian. Variabel penelitian adalah segala sesuatu yang berbentuk apa saja yang ditetapkan oleh peneliti untuk dipelajari sehingga diperoleh informasi tentang hal tersebut, kemudian ditarik kesimpulan (Sudaryono, Gaguk, Rahayu, 2013). Variabel penelitian dibangun oleh teori. Oleh karena itu, pencapaian variabel harus bertumpu pada konstruk teori yang nantinya dikembangkan menjadi definisi operasional pencapaian variabel.

Variabel penelitian terkait erat dengan ranah penilaian. Secara garis besar terdapat 3 ranah penelitian, yaitu kognitif, afektif, dan psikomotorik. Ranah kognitif merupakan kemampuan intelektual yang berkaitan dengan aspek pengetahuan, seperti kemampuan memahami konsep, *Intelligence Quotients (IQ)*, dan kemampuan menganalisis. Ranah afektif merupakan kemampuan seseorang yang berkaitan dengan sikap, seperti minat, tanggung jawab, kejujuran, dan lainnya.

Adapun ranah psikomotorik merupakan kemampuan seseorang yang berkaitan dengan keterampilan atau kemampuan saraf motorik, seperti keterampilan dalam melaksanakan praktikum, keterampilan dalam berbicara, dan lainnya. Pada bahasan selanjutnya hanya akan membatasi penilaian pada ranah afektif saja. Analisis yang tepat terkait variabel penelitian dan ranah penilaian akan menjadi modal yang baik dalam menentukan instrumen penelitian.

B. Penilaian Afektif

Penilaian afektif merupakan penilaian yang menyangkut sikap. Krathwohl (1961) menyatakan bahwa ranah afektif menurut taksonomi terdiri atas 5, yaitu *receiving (attending)*, *responding*, *valuing*, *organisation*, dan *characterization*. Penjelasan masing-masing point akan dikaitkan dengan bidang pendidikan. Pada tahap *receiving* atau *attending*, seseorang memiliki keinginan memperhatikan suatu fenomena khusus atau stimulus, misalnya kelas, kegiatan, musik, buku, dan lainnya. Tugas pendidik dalam hal ini adalah mengarahkan perhatian peserta didik pada fenomena yang menjadi objek pembelajaran afektif.

Responding merupakan partisipasi aktif siswa, yaitu sebagai bagian dari perilakunya. Pada level ini, peserta didik tidak hanya memperhatikan fenomena khusus tetapi juga bereaksi. Hasil pembelajaran pada tahap ini menekankan pada perolehan respons, berkeinginan memberi respons, atau kepuasan dalam memberi respons. Level tertinggi dalam kategori ini adalah minat, yaitu hal-hal yang menekankan pada pencapaian hasil dan kesenangan pada aktivitas khusus seperti kesenangan dalam membaca buku.

Valuing melibatkan penentuan nilai, keyakinan atau sikap dan menunjukkan derajat internalisasi dan komitmen. Rentangan derajat mulai dari menerima suatu nilai, misalnya keinginan untuk meningkatkan keterampilan, sampai pada tingkat komitmen. *Valuing* atau penilaian berbasis pada internalisasi dari seperangkat nilai yang spesifik. Hasil belajar pada level ini berhubungan dengan perilaku yang konsisten dan stabil agar nilai dikenal dengan jelas. Dalam

tujuan pembelajaran, penilaian ini diklasifikasikan sebagai sikap dan apresiasi.

Pada level organisasi, nilai satu dengan nilai lainnya dikaitkan dan konflik antar nilai diselesaikan, dan mulai membangun sistem nilai internal yang konsisten. Hasil pembelajaran pada level ini berupa konsptualisasi nilai atau organisasi sistem nilai. Misalnya, pengembangan filsafat hidup.

Peringkat ranah afektif tertinggi adalah *characterization* nilai. Pada level ini, peserta didik memiliki sistem nilai yang mengendalikan perilaku sampai pada suatu waktu tertentu sehingga terbentuk gaya hidup. Hasil pembelajaran pada level ini berkaitan dengan personal, emosional, dan sosial.

Anderson (1981: 4) menyatakan bahwa pemikiran atau perilaku harus memiliki 3 kriteria untuk diklasifikasikan sebagai ranah afektif, yaitu (1) intensitas, (2) arah, dan (3) target. Intensitas menyatakan derajat atau kekuatan dari perasaan. Beberapa perasaan lebih kuat dari yang lain, misalnya cinta lebih kuat dari senang atau suka. Selain itu, sebagian orang kemungkinan memiliki perasaan yang lebih kuat dibanding yang lain.

Arah berkaitan dengan orientasi positif atau negatif dari suatu perasaan. Arah menunjukkan apakah perasaan itu baik atau buruk. Nilai intensitas dan arah perasaan ditinjau bersama-sama, maka karakteristik afektif berada dalam satu skala yang kontinum.

Target mengacu pada objek, aktivitas atau ide sebagai arah dari perasaan. Jika kecemasan merupakan karakteristik afektif yang ditinjau, ada beberapa kemungkinan target. Peserta didik mungkin bereaksi terhadap sekolah, matematika, situasi sosial, atau pengajaran. Tiap unsur ini bias merupakan target dari kecemasan. Kadang-kadang target ini diketahui oleh seseorang namun kadang tidak diketahui. Seringkali peserta didik merasa tegang bila menghadapi tes di kelas. Peserta didik tersebut cenderung sadar bahwa target ketegangan adalah tes.

McCoach, Gable, dan Madura (2013) menyatakan bahwa afektif merupakan perasaan dan sikap. Instrumen ranah afektif digunakan untuk mengetahui perasaan, sikap, atau emosi seseorang. Ada lima tipe

karakteristik afektif yang penting dalam pelaksanaan pendidikan, yaitu sikap, minat, nilai, konsep diri, dan karakter.

1. Sikap

Sikap yaitu predisposisi yang dipelajari untuk merespon secara positif atau negatif terhadap suatu objek, situasi, konsep, atau orang.

2. Minat

Minat merupakan sikap yang memiliki karakteristik adanya kebutuhan untuk memberi perhatian yang selektif terhadap orang, aktivitas, atau tujuan. Minat yaitu suatu disposisi yang terorganisir melalui pengalaman yang mendorong seseorang untuk memperoleh obyek khusus, aktivitas, pemahaman, dan keterampilan untuk tujuan perhatian atau pencapaian. Hal penting pada minat adalah intensitasnya. Secara umum, minat termasuk karakteristik afektif yang memiliki intensitas tinggi.

3. Nilai

Nilai adalah suatu obyek, aktivitas, atau ide yang dinyatakan oleh individu yang mengendalikan pendidikan dalam mengarahkan minat, sikap, dan kepuasan. Sejak manusia belajar menilai suatu objek, aktivitas dan ide, objek menjadi pengatur penting terhadap minat, sikap, dan kepuasan. Oleh karenanya, sekolah harus menolong peserta didik dalam memperoleh kebahagiaan personal dan memberi kontribusi positif terhadap masyarakat.

Ranah afektif yang tergolong penting dalam proses pembelajaran adalah kejujuran, integritas, adil, kebebasan, dan komitmen. Peserta didik harus belajar untuk menghargai kejujuran dalam berinteraksi dengan orang lain. Peserta didik juga harus mengikat pada kode nilai, misalnya moral, dan artistik. Terhadap nilai adil, peserta didik harus berpendapat bahwa semua orang memperoleh perlakuan hukum yang sama. Terhadap nilai kebebasan, peserta didik harus yakin bahwa negara demokratis harus memberi kebebasan secara maksimum kepada semua orang. Selain itu, peserta didik juga harus menepati janji dan memegang teguh apa yang dikatakan.

4. **Konsep diri**

Konsep diri adalah evaluasi yang dilakukan individu terhadap kemampuan dan kelemahan yang dimilikinya. Target, arah, dan intensitas konsep diri pada dasarnya seperti ranah afektif lainnya. Target konsep diri biasanya orang tetapi juga bisa institusi seperti sekolah. Arah konsep diri dapat positif tetapi juga dapat negatif. Intensitasnya dapat dinyatakan dalam suatu daerah kontinum, yaitu mulai dari yang rendah sampai yang tertinggi. Konsep diri ini penting untuk menentukan jenjang karir peserta didik, yaitu dengan mengetahui kekuatan dan kelemahan diri sendiri sehingga bisa memilih karir yang tepat bagi diri sendiri. Selain itu, informasi konsep diri ini penting bagi sekolah untuk memotivasi belajar peserta didik.

5. **Karakter**

Karakter merupakan bagian ranah afektif yang berbeda dengan ranah kognitif tetapi ada kaitannya. Karakter merupakan kualitas moral dan yang menjadi pertimbangan seseorang dalam membuat keputusan, bersikap, dan berperilaku yang bersifat relatif tetap. Kant (Aune: 1979) menyatakan bahwa berperilaku secara moral adalah suatu kewajiban untuk memenuhi aturan yang berlaku di masyarakat. Moral diartikan sebagai perilaku yang sesuai dengan standar baik dan benar atau penetapan baik dan buruk perilaku atau karakter seseorang. Jadi, karakter dapat dinilai dari perilaku yang ditampilkan di masyarakat seseorang berada.

Semua penilaian terhadap data penelitian membutuhkan instrumen. Adapun jenis instrumen yang diperlukan sangat bergantung dengan variabel yang akan dicapai. Secara garis besar, instrumen penelitian dibedakan atas 2 jenis, yaitu tes dan nontes. Tes merupakan instrumen untuk mengukur karakteristik seseorang atau sekelompok orang terkait kemampuan kognitif atau keterampilan seseorang. Adapun instrumen nontes adalah instrumen untuk mengukur karakteristik seseorang atau sekelompok orang terkait kemampuan afektif seperti tindakan, sikap, perilaku, etika, dan sebagainya. Instrumen nontes seperti: lembar angket, lembar observasi, dan lembar pedoman

wawancara. Lembar angket merupakan instrumen penelitian yang di dalamnya memuat beberapa pernyataan yang harus diisi terkait sikap responden. Lembar observasi memuat penilaian peneliti terhadap sikap atau fenomena yang teramati. Lembar wawancara memuat beberapa pertanyaan yang akan digunakan untuk menggali informasi responden. Berdasarkan jenis instrumennya, penilaian afektif memerlukan instrumen nontes. Pada bahasan selanjutnya hanya akan membahas penilaian afektif menggunakan instrumen angket.

C. Angket

Angket atau kuesioner merupakan teknik atau cara pengumpulan data secara tidak langsung (penelitian tidak langsung bertanya-jawab dengan responden). Instrumen atau alat pengumpulan datanya juga disebut angket berisi sejumlah pertanyaan atau pernyataan yang harus dijawab atau direspon oleh responden. Bentuk pertanyaan pada angket dapat bermacam-macam, yaitu pertanyaan terbuka, pertanyaan terstruktur dan pertanyaan tertutup (Cresswell, 2007). Artinya, angket adalah daftar pertanyaan yang diberikan kepada orang lain bersedia memberikan respon sesuai dengan permintaan pengguna.

Tujuan penyebaran angket ialah mencari informasi yang lengkap mengenai suatu masalah dari responden tanpa merasa khawatir bila responden memberikan jawaban yang tidak sesuai dengan kenyataan dalam pengisian daftar pertanyaan. Di samping itu, responden mengetahui informasi tertentu yang diminta.

Angket cocok digunakan bila jumlah responden cukup besar dan tersebar di wilayah yang cukup luas. Angket dapat diberikan kepada responden secara langsung atau dikirim melalui pos atau internet.

Berdasarkan siapa yang menjawab, angket dibedakan atas 2, yaitu angket langsung dan tidak langsung. Angket langsung berarti bahwa angket dikirimkan dan diisi langsung oleh orang yang akan dimintai jawaban tentang dirinya. Angket tidak langsung yaitu angket yang dikirimkan dan diisi oleh bukan orang yang diminta keterangannya. Kuesioner tidak langsung biasanya digunakan untuk mencari informasi tentang bawahan, anak, saudara, tetangga, dan sebagainya.

Dilihat dari segi cara menjawabnya, angket dibedakan menjadi 2 yaitu angket terbuka dan angket tertutup.

1. Angket terbuka

Angket terbuka berisi pernyataan-pernyataan atau pertanyaan pokok yang bisa dijawab atau direspon oleh responden secara bebas. Tidak ada anak pertanyaan rincian yang memberikan arah dalam pemberian jawaban atau respon. Responden mempunyai kebebasan untuk memberikan jawaban atau respon sesuai dengan persepsinya. Dengan kata lain yang dimaksud dengan angket terbuka (angket tidak terstruktur) adalah angket yang disajikan dalam bentuk sederhana sehingga responden dapat memberikan isian sesuai dengan kehendak dan keadaanya.

Angket terbuka memiliki keuntungan baik bagi responden maupun peneliti. Responden dapat mengisi sesuai dengan keinginan yang sesuai dengan keadaan yang dialaminya. Adapun bagi peneliti akan mendapat data yang bervariasi, bukan hanya yang sudah disajikan karena sudah diasumsikan oleh peneliti.

Contoh 2.1

Pelatihan apa saja yang pernah saudara ikuti? Tuliskan sebenarnya, dimana dan tahun berapa lulusnya.

No	Pelatihan	Tempat	Tahun lulus
1
2
3
dst

Contoh 2.2

Bagaimanakah pendapat saudara tentang pembelajaran daring? Apakah saudara pernah mengalaminya? Jika pernah, bagaimana komentar Anda?

2. Angket tertutup

Pada angket tertutup, pernyataan atau pertanyaan yang sudah disusun secara terstruktur di samping ada pertanyaan pokok atau pertanyaan utama, juga ada anak pertanyaan atau sub pertanyaan. Dalam angket tertutup, pertanyaan-pertanyaan telah memiliki

alternatif jawaban yang tinggal dipilih oleh responden. Artinya, angket terstruktur adalah angket yang disajikan dalam bentuk sedemikian rupa sehingga responden diminta untuk memilih satu jawaban yang sesuai dengan karakteristik dirinya dengan cara memberikan tanda silang (X) atau ceklist (√). Responden tidak dapat memberikan jawaban atau responden lain kecuali yang telah tersedia sebagai alternatif jawaban (Djaali, 2004).

Contoh 2.3

Cara memberi tanda silang (x)

Apakah saudara pernah mempraktekkan metode pembelajaran *active learning*?

(1) Pernah

(2) Tidak pernah

Jika pernah, strategi apa saja yang Anda gunakan?

- a. Pembelajaran *authentic*
- b. Pembelajaran berbasis inkuiri
- c. Pembelajaran berbasis masalah
- d. Pembelajaran berbasis kerja

Contoh 2.4

Apakah Anda merupakan orang yang aktif membaca buku?

(1) Ya

(2) Tidak

Jika ya, berapa lama Anda membaca dalam satu hari?

- a. 0 menit – 30 menit
- b. 31 menit – 60 menit
- c. 61 menit – 90 menit
- d. 91 menit – 120 menit

Contoh 2.5

Apakah Anda merupakan mahasiswa yang aktif dalam kegiatan pembelajaran?

(1) Ya

(2) tidak

Jika ya, berapa rata-rata anda bertanya atau memberi tanggapan atas pertanyaan guru di kelas dalam setiap satu pembelajaran?

- a. 1 kali
- b. 2 kali
- c. 3 kali
- d. 4 kali

Angket menjadi suatu instrumen yang diisi secara bebas oleh responden. Agar peneliti mendapatkan jawaban responden yang valid dan benar-benar mencuplik keadaan responden, maka ada beberapa hal yang wajib diperhatikan dalam menyusun angket.

1. Perlu adanya pengantar dan petunjuk pengisian angket. Pada pengantar terdapat tujuan pengedaran angket, jaminan kerahasiaan jawaban serta ucapan terima kasih kepada responden. Petunjuk pengisian menjelaskan bagaimana cara menjawab pertanyaan atau merespon pertanyaan tersebut.
2. Butir-butir pertanyaan dirumuskan secara jelas, menggunakan kata-kata yang lazim digunakan (populer), kalimat tidak terlalu panjang dan tidak beranak pinak. Pada butir-butir pertanyaan atau pernyataan tertutup sebaiknya hanya berisi satu pesan, sedangkan dalam pertanyaan atau pernyataan terbuka bisa berisi satu pesan kompleks atau lebih dari satu pesan yang tidak terlalu kompleks. Pada angket tertutup sebaiknya hanya bersis satu pesan yang tidak terlalu kompleks.
3. Untuk setiap pernyataan atau pertanyaan terbuka dan terstruktur disediakan kolom untuk menuliskan jawaban atau respon dari responden secukupnya. Untuk pertanyaan atau pernyataan tertutup telah disediakan alternatif jawaban dan tiap alternatif hanya berisi satu pesan sederhana. Jawaban atau respon dari responden dapat langsung diberikan pada alternatif jawaban, atau menggunakan lembar jawaban khusus bersatu atau terpisah dari lembar pertanyaan atau pernyataan.

Terkait penyusunan pernyataan dalam angket, Mardapi (2008) menyarankan untuk memperhatikan beberapa hal berikut, yaitu:

1. Gunakan kata-kata yang sederhana sesuai dengan tingkat pendidikan;
2. Tidak menggunakan pertanyaan yang bersifat samar-samar
3. Hindari pertanyaan yang bias;
4. Hindari pertanyaan yang bersifat penganandaian;
5. Jumlah pertanyaan perlu diperhatikan agar tidak menimbulkan kebosanan responden;
6. Lama pengisian instrumen sebaiknya tidak lebih dari 20 menit.

Kelebihan dan kelemahan angket

Usman dan Akbar (2003: 72) menyatakan keuntungan dan kelemahan menggunakan instrumen angket. Keuntungan menggunakan angket tertutup yaitu (1) mudah diolah, (2) responden tidak perlu menuliskan buah pikirannya, (3) pengisian menggunakan waktu yang singkat, (4) dapat menjangkau responden yang relative banyak karena responden lebih mendalam. Adapun keuntungan angket terbuka antara lain: (1) responden dapat mengungkapkan buah pikirannya, (2) berguna bila peneliti ingin mengetahui keadaan responden lebih mendalam. Kelemahan menggunakan angket tertutup, yaitu: (1) responden tidak memiliki kesempatan untuk menjawab lebih bebas dan (2) ada kemungkinan responden asal mengisi saja. Adapun kelemahan menggunakan angket terbuka, meliputi: (1) sukar mengolahnya, (2) perlu waktu yang relative panjang untuk mengisinya, dan (3) nilai jawaban yang tidak sama.

Terlepas adanya kelebihan dan kekurangan angket, semua instrumen angket harus dilengkapi dengan skala yang berfungsi untuk mengukur variabel penelitian. Banyaknya jenis skala penilaian perlu dicermati dengan baik guna menentukan skala yang tepat guna mempermudah baik peneliti mendapatkan data dan responden dalam menggunakan instrument. Bahasan berikut akan mengkaji lebih jauh terkait skala penilaian.

D. Skala Penilaian

Skala adalah alat yang disusun dan digunakan oleh peneliti untuk mengubah respon tentang suatu variabel yang bersifat kualitatif menjadi kuantitatif (Mahmud, 2011). Dalam kegiatan pengukuran, variabel yang bersifat kualitatif berskala nominal, sedangkan variabel kuantitatif berskala ordinal, interval, atau rasio. Melalui pengubahan ini, variabel berskala nominal diubah ke dalam variabel berskala interval. Jadi, dalam konteks penelitian, penggunaan instrumen skala dimaksudkan untuk menghimpun data berskala interval (Muhammad Ali, 2009).

Hasil dari skala harus diinterpretasikan secara hati-hati karena selain menghasilkan gambaran yang kasar, jawaban responden tidak begitu saja langsung mudah dipercaya. Terdapat beberapa aspek yang menjadikan jawaban responden tidak dapat mudah dipercaya, antara lain persahabatan, kecepatan menerka, kecepatan dalam memutuskan, jawaban kesan pertama, penampilan instrumen, prasangka, dan kemurahan hati (Arikunto, 2010). Penggunaan instrumen skala ditujukan pada pengumpulan data yang berhubungan dengan aspek emosional objek penelitian. Lebih lanjut Muhammad Ali dalam Mahmud (2011: 182) mengemukakan bahwa data tentang sikap, motivasi, minat, dan penilaian dikumpulkan melalui instrumen skala.

Sikap adalah kecenderungan tentang perilaku seseorang terhadap suatu objek, orang, atau perilaku orang lain. Kecenderungan ini ditunjukkan dengan derajat kesetujuan dan ketidaksetujuannya terhadap sesuatu yang menjadi sasaran kecenderungan tersebut. Motivasi adalah derajat dorongan yang ada dalam frekuensi atau keseringan, kesungguhan atau ketekunan, dan lamanya seseorang bertahan melakukan suatu kegiatan untuk mencapai tujuan tertentu. Minat adalah derajat preferensi pilihan suka atau tidak suka terhadap suatu objek atau kegiatan ditimbulkan ketertarikan orang tersebut pada objek atau kegiatan tersebut. Penilaian adalah derajat kualitas yang dinilai berdasarkan pandangan seseorang terhadap suatu objek, kegiatan, atau orang lain.

Salah satu instrumen yang menggunakan konsep skala adalah angket. Pengubahan respon angket dalam bentuk skala menurut

Sudarmono, dkk: 2013 memudahkan peneliti untuk menguji validitas dan reliabilitas instrumen tersebut, terutama apabila pengujiannya akan dilakukan melalui analisis statistik.

Macam-Macam Skala Pengukuran

Skala pengukuran adalah kesepakatan yang digunakan sebagai acuan untuk menentukan panjang pendeknya interval yang ada dalam alat ukur, sehingga alat ukur tersebut bila digunakan dalam pengukuran akan menghasilkan data kuantitatif. Data statistik berkaitan erat dengan hasil skala pengukuran yang digunakan. Skala pengukuran ada 4, yaitu skala nominal, skala ordinal, skala interval, dan skala ratio. Skala pengukuran ini akan menghasilkan data nominal, data ordinal, data interval, dan data rasio.

Skala nominal adalah skala yang paling sederhana disusun menurut jenis (kategorinya) atau fungsi bilangan hanya sebagai simbol untuk membedakan sebuah karakteristik dengan karakteristik lainnya. Misalnya, jenis mata pencaharian dibedakan atas: 1 = PNS, 2 = pedagang, 3 = buruh, 4 = petani.

Skala ordinal adalah skala yang didasarkan pada ranking, diurutkan dari jenjang yang lebih tinggi sampai jenjang terendah atau sebaliknya. Skala ordinal memungkinkan untuk mengurutkan seseorang atau objek sesuai dengan banyak atau kuantitas dari karakteristik yang dimilikinya. Skala ordinal, memungkinkan untuk melakukan perhitungan atau kuantifikasi variable-variabel yang diuji sehingga dapat memberikan informasi yang lebih substansial dibanding dengan skala nominal. Misalnya, mengukur kejuaraan, yaitu juara 1, juara 2, dan juara 3. Contoh lainnya yaitu status sosial (kaya, sederhana, dan miskin).

Skala interval adalah skala yang menunjukkan jarak antara satu data dengan data yang lain dan mempunyai bobot yang sama. Skala pengukuran interval merupakan angka untuk operasi aritmatika. Skala interval hampir sama dengan skala ordinal. Hanya saja, pada skala interval memiliki tambahan karakteristik yaitu interval yang tetap. Misalnya, skor IQ, waktu (menit, jam, hari, bulan, tahun).

Skala rasio adalah skala pengukuran yang mempunyai nol mutlak dan jarak yang sama. Dengan kata lain, data berskala rasio adalah data yang diperoleh dengan cara pengukuran di mana jarak dua titik pada skala sudah diketahui. Contoh: umur, tinggi badan, dan berat badan manusia (Sudaryono, Margono, & Rahayu, 2013).

Selain pembagian berdasarkan jenis datanya, para ahli sosiologi membedakan skala dalam dua tipe pengukuran menurut gejala sosial yang diukur.

1. Skala pengukuran untuk mengukur perilaku susila dan kepribadian. Contohnya, yaitu skala sikap, skala moral, tes karakter, dan skala partisipasi social.
2. Skala pengukuran untuk mengukur berbagai aspek budaya lain dan lingkungan social, seperti skala mengukur status sosial ekonomi, lembaga-lembaga swadaya masyarakat (sosial), kemasyarakatan, kondisi rumah tangga, dan lain sebagainya.

Pada pembahasan kali ini, hanya akan dibahas tentang skala sikap. Ada 7 skala sikap yang sering digunakan di berbagai penelitian, yaitu: skala Likert, skala Guttman, skala perbedaan semantik (*semantic differential*), skala penilaian (*rating scale*), skala Borgadus, skala sosiometrik, dan skala Thurstone.

1. Skala Likert

Skala Likert digunakan untuk mengukur sikap, pendapat, dan persepsi seseorang atau sekelompok tentang kejadian atau gejala sosial (Sudaryono, Margono, & Rahayu, 2013). Dalam penelitian gejala sosial telah ditetapkan secara spesifik oleh peneliti, yang selanjutnya disebut sebagai variabel penelitian. Dengan menggunakan skala Likert, maka variabel yang akan diukur dijabarkan menjadi dimensi, dimensi dijabarkan menjadi sub-variabel, kemudian sub-variabel dijabarkan lagi menjadi indikator-indikator yang dapat diukur. Akhirnya, indikator tersebut dapat dijadikan sebagai titik tolak untuk membuat item instrumen berupa pertanyaan atau pernyataan. Setiap jawaban

yang dihubungkan dengan bentuk pernyataan atau dukungan sikap yang diungkapkan dengan kata-kata sebagai berikut.

*Sangat tinggi/sangat penting/sangat benar/sangat setuju sekali
Tinggi/penting/benar/setuju*

Cukup tinggi/cukup penting/ cukup benar/netral

Rendah/tidak penting/salah/tidak setuju

Rendah sekali/sangat tidak penting/sangat salah/sangat tidak setuju

Dalam penyusunan instrumen untuk variabel tertentu, sebaiknya butir-butir pernyataan dibuat dalam bentuk kalimat positif, netral, atau negatif sehingga responden dapat menjawab serius dan konsisten.

Instrumen penelitian yang menggunakan skala Likert dapat dibuat dalam bentuk checklist ataupun pilihan ganda.

Contoh 2.6

(Skala Likert bentuk checklist)

Berilah jawaban pertanyaan berikut sesuai pendapat saudara dengan memberi tanda (√) dalam kolom yang tersedia.

No	Pertanyaan	Jawaban				
		SS	ST	RG	TS	STS
1.	Sekolah ini akan menggunakan teknologi informasi dalam pelayanan administrasi dan akademik		√			

Keterangan:

- SS = Sangat Setuju 5
- ST = Setuju 4
- RG = Ragu-ragu 3
- TS = Tidak Setuju 2
- STS = Sangat Tidak Setuju 1

Contoh 2.7

(Skala Likert bentuk pilihan ganda)

Angket untuk mengukur kesempatan yang sama bagi wanita

Pertanyaan

1. Perempuan memiliki lebih sedikit kesempatan dibanding laki-laki untuk mendapatkan pendidikan sebagai bekal mendapatkan pekerjaan.
 - A. Sangat setuju
 - B. Agak setuju
 - C. Tidak tahu
 - D. Tidak setuju
 - E. Sangat Tidak Setuju
2. Banyak perempuan berkualitas tidak bisa mendapatkan pekerjaan yang lebih baik dibanding pria.
 - A. Sangat setuju
 - B. Agak setuju
 - C. Tidak tahu
 - D. Tidak setuju
 - E. Sangat Tidak Setuju
3. Masyarakat kita mendiskriminasi perempuan.
 - A. Sangat setuju
 - B. Agak setuju
 - C. Tidak tahu
 - D. Tidak setuju
 - E. Sangat Tidak Setuju

Penentuan skor

5 = sangat setuju

4 = setuju

3 = Tidak tahu

2 = tidak setuju

1 = sangat tidak tahu

Skor tertinggi yang mungkin = 15

Skor terendah yang mungkin = 3

Contoh 2.8

Angket untuk mengukur percaya diri (*self-esteem*)

Pertanyaan

1. Secara keseluruhan, saya puas dengan diri saya sendiri
 - A. Setuju
 - B. Tidak tahu
 - C. Tidak setuju
2. Kadang-kadang, saya pikir saya sama sekali tidak baik.
 - A. Setuju
 - B. Tidak tahu
 - C. Tidak setuju
3. Saya terkadang merasa bahwa (orang lain) tidak menganggap serius pendapat saya.
 - A. Setuju
 - B. Tidak tahu
 - C. Tidak setuju

Penentuan skor

3 = setuju

2 = tidak tahu

1 = tidak setuju

Skor tertinggi yang mungkin = 9

Skor terendah yang mungkin = 3

2. Skala Guttman

Skala guttman merupakan skala kumulatif. Jika seseorang menyisakan pertanyaan yang berbobot lebih berat, ia akan mengiyakan pertanyaan yang berbobot lainnya. Hal yang sama juga disebutkan oleh Neuman (2014). Skala guttman mengukur suatu dimensi saja dari suatu variabel yang multidimensi. Skala Guttman disebut juga skala skalogram yang sangat baik untuk meyakinkan peneliti tentang kesatuan dimensi dari sikap atau sifat yang diteliti, yang sering disebut dengan atribut universal. Pada skal Guttman terdapat beberapa pertanyaan yang diurutkan secara hirarkis untuk

melihat sikap tertentu seseorang. Jika seseorang menyatakan tidak terhadap pertanyaan sikap tertentu dari sederetan pertanyaan itu, ia akan menyatakan lebih dari tidak terhadap pernyataan berikutnya.

Ada 2 kelemahan pokok skala Guttman, yaitu (1) skala guttman bisa tidak mungkin menjadi dasar yang efektif baik untuk mengukur sikap terhadap objek yang kompleks ataupun untuk membuat prediksi tentang perilaku objek tersebut, (2) satu skala bisa saja mempunyai dimensi tunggal untuk satu kelompok tetapi ganda untuk kelompok lain, ataupun berdimensi satu untuk satu waktu dan mempunyai dimensi ganda untuk waktu yang lain dengan responden yang sama. Jadi, skala guttman merupakan skala yang digunakan untuk yang menjawab yang bersifat jelas (tegas) dan konsisten.

Contoh 2.9

Apa komentar saudara, jika koruptor dihukum mati?

- a. Setuju
- b. Tidak setuju

Saudara sudah menikah?

- a. Sudah
- b. Belum

Jawaban responden dapat diganti dengan nilai (1) dan (0). Analisis data menggunakan skala guttman sama seperti saat menggunakan skala Likert.

3. Skala diferensial semantik (*Semantic Differential Scale*)

Nama lain skala diferensial semantik adalah skala perbedaan semantik. Skala ini berisikan serangkaian karakteristik bipolar (dua kutub), seperti panas-dingin, baik – tidak baik, dan sebagainya. Karakteristik bipolar mempunyai 3 dimensi dasar sikap seseorang terhadap objek, yaitu (1) potensi, yaitu kekuatan atau atraksi fisik suatu objek, (2) evaluasi, yaitu hal-hal yang menguntungkan atau tidak menguntungkan suatu objek, dan (3) aktivitas, yaitu tingkat gerakan suatu objek.

Contoh 2.10

Responden diminta menilai tingkat profesionalitas seseorang dalam mengajar dengan cara melingkari terhadap skor yang tersedia.

Tidak profesional	1	2	3	4	5	Profesional
--------------------------	----------	----------	----------	----------	----------	--------------------

Contoh 2.11

Apakah Anda suka memancing? (pilih salah satu)

.....
Sangat tidak suka	Tidak suka	netral	suka	Sangat suka
(=1)	(=2)	(=3)	(=4)	(=5)

Contoh 2.12

Petunjuk

Silakan baca setiap pasangan kata sifat di bawah ini. Kemudian, beri tanda pada ruang kosong yang paling dekat dengan perasaan kesan pertama Anda dengan tanda centang (✓). Tidak ada jawaban benar atau salah.

Pertanyaan

Apa pendapat anda tentang plagiasi (mencontek)

Buruk	___	___	___	___	___	___	___	Bagus
Adil	___	___	___	___	___	___	___	Tidak adil
Tersembunyi	___	___	___	___	___	___	___	Jelas
Modern	___	___	___	___	___	___	___	Tradisional
Simpel	___	___	___	___	___	___	___	Kompleks
Tersembunyi	___	___	___	___	___	___	___	Jelas
Kotor	___	___	___	___	___	___	___	Bersih

4. Rating skala (skala penilaian)

Pada skala penilaian, si pemberi angka berada pada suatu rangkaian di mana individu atau objek akan ditempatkan. Penilai terdiri dari beberapa orang dan penilai adalah orang yang mengetahui bidang yang dinilai. Pada model ini, responden tidak akan menjawab dari data kualitatif yang sudah tersedia, tetapi menjawab salah satu dari jawaban kuantitatif yang telah disediakan. Jadi, bentuk rating scale lebih fleksibel, tidak terbatas untuk pengukuran sikap saja, tetapi

untuk mengukur persepsi responden terhadap gejala atau fenomena lainnya. Misalnya skala untuk mengukur status sosial ekonomi, IPTEK, instansi dan lembaga, kinerja dosen, kegiatan PBM, kepuasan pelanggan, dan produktivitas kerja, motivasi pegawai, dan lainnya (Djaali, 2004).

Contoh 2.13

Survey terkait tata ruang kelas di SD X

No	Pernyataan	Interval Jawaban			
		4	3	2	1
1	Pencahayaan di setiap ruangan				
2	Kebersihan di setiap ruangan				
3	Ventilasi di setiap ruangan				
4	Penataan meja pendidik dan peserta didik guna mempermudah interaksi				

Penerapan Isu Lingkungan 2.1

Penelitian bertujuan untuk menganalisis pengaruh motivasi belajar terhadap hasil belajar siswa. Instrumen yang digunakan angket menggunakan rating scale. Bagaimana penggunaan rating scale pada angket yang disusun, silahkan langsung cek di link <https://ejournal.upi.edu/index.php/jpmanper/article/view/14958/8522>. Sumber: Andriani, N & Rasto. (2019). Motivasi belajar sebagai determinan hasil belajar siswa. *Jurnal Pendidikan Manajemen Perkantoran*, 4(1): 80-86. Atau melalui link berikut.

5. Skala Thurstone

Skala Thurstone meminta responden untuk memilih pertanyaan yang ia setuju dari beberapa pertanyaan yang menyajikan pandangan yang berbeda-beda. Pada umumnya setiap item mempunyai asosiasi nilai antara 1 sampai 10, tetapi nilai-nilainya tidak diketahui oleh responden. Pemberian nilai ini berdasarkan jumlah tertentu pertanyaan yang dipilih oleh responden mengenai angket tersebut (Subana, 2004). Perbedaan antara skala Thurstone dan Likert adalah pada skala

thurstone yang panjangnya sama memiliki intensitas kekuatan yang sama, sedangkan pada skala Likert tidak perlu sama. Pada skala Thurstone rentang skala yang disediakan lebih dari lima pilihan dan disarankan sekitar sepuluh pilihan jawaban (misalnya dengan rentang angka 1 s/d 11 atau a s/d k). Jawaban di tengah adalah netral, semakin ke kiri semakin tidak setuju, sebaliknya semakin ke kanan semakin setuju.

1	2	3	4	5	6	7	8	9	10	11
---	---	---	---	---	---	---	---	---	----	----

Tidak Setuju

Setuju

Demikian berbagai skala yang dapat digunakan dalam penelitian. Ketepatan dalam menentukan skala sangat diperlukan guna mencapai hasil penelitian yang baik. Cara menentukan skala yang baik juga menjadi bagian dalam langkah-langkah pengembangan instrumen yang akan dibahas secara lebih detail pada BAB III.

BAB III

PENGEMBANGAN INSTRUMEN

A. Langkah Pengembangan Instrumen

Proses pengumpulan data merupakan suatu hal yang sangat penting dalam suatu penelitian. Data hasil penelitian sangat bermanfaat untuk menjawab tujuan penelitian dan menguji hipotesis (Wiersma, 1986). Sebagai contoh misalnya pada penelitian eksploratif, hasil pengumpulan data digunakan untuk penyimpulan dengan membuat deskripsi untuk mengeksplorasi hal-hal terkait permasalahan penelitian. Pada penelitian positivistik, hasil pengumpulan data dianalisis dengan uji statistik tertentu, hasil analisis digunakan untuk membuat kesimpulan.

Proses pengumpulan data tidak dapat dilepaskan dengan instrumen. Instrumen yang baik akan mampu menghasilkan data yang baik. Data yang baik akan berefek pada hasil penelitian yang baik pula. Oleh karena itu, instrumen perlu disusun dengan baik guna menghasilkan instrumen yang memenuhi standar baku.

Suryabrata (2005) menyebutkan 9 (sembilan) langkah pengembangan instrumen non tes yang baik, yaitu (1) pengembangan spesifikasi alat ukur, (2) penulisan pernyataan dan pertanyaan, (3) penelaahan pernyataan atau pertanyaan, (4) perakitan instrument (untuk keperluan ujicoba), (5) ujicoba, (6) analisis hasil ujicoba, (7) seleksi dan perakitan instrument, (8) administrasi instrument (bentuk akhir), dan (9) penyusunan skala dan norma. Mardapi (2008) menyatakan 10 langkah dalam mengembangkan instrumen afektif, yaitu (1) menentukan spesifikasi instrumen, (2) menulis instrumen, (3) menentukan skala instrumen, (4) menentukan sistem penskoran, (5) mentelaah instrumen, (6) melakukan ujicoba, (7) menganalisis instrumen, (8) merakit instrumen, (9) melaksanakan pengukuran, dan (10) menafsirkan hasil pengukuran. Berikut penjelasan masing-masing langkah tersebut.

1. Menentukan spesifikasi instrumen

Penentuan spesifikasi instrumen merupakan langkah utama dalam mengembangkan instrumen penilaian berupa penentuan tujuan dan kisi-kisi instrumen. Menurut Mardapi (2010), ada 4 hal yang harus diperhatikan dalam menyusun spesifikasi instrumen

a. Menentukan tujuan penilaian

Tujuan penelitian menjadi fokus utama pengembangan instrumen. Semakin jelas tujuan penelitian akan semakin mudah dalam menentukan langkah penelitian lainnya. Contoh tujuan penelitian dalam koridor pengembangan instrumen adalah mengembangkan instrumen untuk mengukur sikap kerja sinergis siswa dalam melaksanakan pembelajaran.

b. Menyusun kisi-kisi instrumen

Istilah lain dari kisi-kisi instrumen adalah *blue-print*. Kisi-kisi berwujud tabel matrik yang berisi spesifikasi instrument yang akan ditulis. Tahapan penyusunan kisi-kisi instrumen meliputi (1) menentukan definisi konseptual suatu variabel penelitian yang diambil dari teori, (2) penjabaran definisi operasional menjadi sejumlah indikator instrumen. Indikator inilah acuan yang penyusunan instrumen.

c. Memilih bentuk dan format instrumen

Format instrumen terkait dengan jenis data. Jenis data yang dimaksud dapat berupa data nominal, data ordinal, data interval dan data rasio. Data nominal merupakan ukuran diskrit (terpisah antar data) yang menunjukkan karakteristik individu. Contoh data nominal seperti agama, warna pakaian, kendaraan, jenis kelamin, dan hobi. Data ordinal merupakan ukuran yang menunjukkan posisi suatu objek, dengan ukuran tersebut dapat diurutkan dari urutan paling rendah sampai yang paling tinggi, namun belum ada jarak atau interval antara posisi ukuran yang satu dengan yang lain. Misalnya skala Likert (sangat setuju, setuju, ragu-ragu, tidak setuju, sangat tidak setuju), dimana belum ada jarak yang jelas antara tidak setuju dengan sangat tidak setuju, dan juga skala lainnya. Data

interval merupakan ukuran yang menunjukkan posisi suatu objek dalam suatu urutan paling rendah sampai paling tinggi, dan ada jarak atau interval antara posisi ukuran yang satu dengan lain. Contoh data ini adalah nilai/skor dalam pendidikan. Pada data interval, nilai nol juga bukan nilai yang mutlak, yang berarti bahwa seorang peserta didik memperoleh skor nol, belum tentu peserta didik tersebut sama sekali tidak menguasai kompetensi dalam pembelajaran, namun bisa jadi karena alasan lain. Pada data rasio, ukuran menunjukkan posisi suatu objek dalam suatu skala paling rendah sampai skala yang paling tinggi, ada jarak atau interval antara posisi ukuran yang satu dengan yang lain, dan adanya besaran absolute/mutlak. Sebagai contoh pada data rasio adalah ukuran volume air. Volume air dalam suatu wadah sama dengan nol berarti air dalam wadah tersebut memang telah kosong, atau tidak ada air sedikitpun dalam wadah tersebut.

Penentuan jenis data berdampak pada pelaksanaan pengukuran dalam penelitian. Sebagai contoh seorang peneliti ingin mengetahui motivasi kerja karyawan. Permasalahan yang menjadi fokus penelitian adalah motivasi, yang dapat diukur dengan angket/kuisisioner motivasi. Untuk pengumpulan data ini, perlu digunakan teknik nontes. Teknik nontes selain angket, yaitu wawancara, observasi, dan dokumentasi.

d. Menentukan panjang instrumen.

Panjang instrumen terkait dengan jumlah item instrumen dan derajat kedalaman instrumen dalam menggali informasi responden. Untuk tujuan penelitian yang sama, semakin sedikit jumlah item instrumen maka pertanyaan instrumen semakin detail. Sebaliknya, semakin banyak jumlah item instrumen maka pertanyaan item instrument semakin singkat.

Ada bermacam-macam teknik penyusunan pertanyaan, di antaranya pertanyaan dikotomi, pertanyaan pilihan ganda, urutan bertingkat (*rank ordering*), *rating scale*, dan pertanyaan terbuka yang masing-masing bentuk memiliki ciri

khas tersendiri. Penjelasan masing-masing teknik penyusunan pertanyaan dapat dilihat kembali pada BAB II.

2. Menulis instrumen

Penulisan instrumen dimulai dengan penentuan definisi konseptual dan definisi operasional. Definisi konseptual adalah definisi sebuah variabel yang didasarkan pada teori-teori yang sudah ada. Definisi operasional merupakan definisi yang disusun berdasarkan bukti yang dapat terdeteksi dari data lapangan. Definisi operasional kemudian dijabarkan menjadi sejumlah indikator. Indikator ini kemudian dijadikan sebagai pedoman penyusunan item untuk kemudian digunakan dalam penyusunan instrument.

Sudaryono, Margono, dan Rahayu (2013) menyarankan bagi peneliti yang akan menggunakan angket sebagai instrument penelitian seyogyanya memperhatikan ketentuan atau kaidah penulisannya meliputi 3 hal, yaitu isi (materi), bahasa, dan konstruksi.

a. Isi (materi)

Pernyataan harus sesuai dengan rumusan indikator dalam kisi-kisi. Selain itu, aspek yang diukur pada setiap pernyataan sudah sesuai dengan tuntutan dalam kisi-kisi.

b. Bahasa

Bahasa harus komunikatif dan sesuai dengan jenjang pendidikan responden. Selain itu, akan lebih baik jika menggunakan bahasa Indonesia baku dan tidak menggunakan bahasa yang tabu.

c. Konstruksi

Pernyataan sebaiknya dirumuskan secara singkat dan jelas. Kalimat bebas dari pernyataan yang tidak relevan dengan objek yang dipersoalkan. Kalimat juga diharapkan bebas dari pernyataan yang bersifat negatif. Kalimat bebas dari pernyataan yang mengacu pada masa lalu. Kalimat bebas dari pernyataan yang dapat diinterpretasikan lebih dari satu cara. Kalimat bebas dari pernyataan yang mungkin disetujui atau

dikosongkan oleh hampir semua responden. Setiap pernyataan hanya berisi satu gagasan secara lengkap. Kalimat juga sebaiknya bebas dari pernyataan yang tidak pasti, seperti semua, selalu, kadang-kadang, tidak satu pun, dan tidak pernah. Selain itu, kata-kata hanya, sekedar, sama sekali, dan /atau semata-mata harus digunakan seperlunya dan dengan hati-hati agar tidak menimbulkan kesalahan penafsiran.

3. Menentukan skala pengukuran

Pada umumnya, skala instrumen yang sering digunakan dalam penelitian, yaitu skala Likert, skala Thurstone, dan skala beda semantik. Skala Likert merupakan skala pengukuran yang paling sederhana dan mudah dibuat, yaitu dengan cara:

- a. menetapkan tujuan pengukuran
- b. menentukan definisi konseptual
- c. menyusun definisi operasional
- d. menentukan indikator
- e. menentukan alat ukur
- f. menyusun item berdasarkan indikator yang telah ditentukan.

Setelah itu, item diujicobakan ke responden. Data hasil ujicoba kemudian dianalisis untuk mengetahui validitas dan reliabilitasnya. Data pada skala Likert berjenis ordinal. Data pada skala Thurstone berjenis interval. Penyusunan skala Thurstone sedikit lebih rumit disbanding skala Likert. Meskipun demikian, ada bagian yang sama dengan skala Likert, seperti:

- a. Penetapan tujuan pengukuran
- b. Menentukan definisi konseptual
- c. Menyusun definisi operasional
- d. Menentukan indikator
- e. Menentukan alat ukur
- f. Menyusun item berdasarkan indikator yang telah ditentukan.

Jadi, pembeda antara instrumen menggunakan skala Likert dan skala Thurstone terletak pada perlakuan setelah item tersusun. Langkah tersebut yaitu membuat format untuk proses penilaian yang dilakukan oleh ahli. Setiap item diberikan alternatif respon, biasanya dengan rentang 11 atau 7. Langkah selanjutnya yaitu mencari beberapa ahli untuk memberikan penilaian pada item yang telah tersusun. Petunjuk yang diberikan kepada ahli yaitu menilai masing-masing item dengan meletakkan pada point tertentu. Semakin mendekati point tinggi maka item semakin *favorable*, sebaliknya semakin rendah maka semakin *unfavorable*. Setelah proses penilaian selesai, maka langkah selanjutnya yaitu melakukan tabulasi data dan pemilihan item. Item terpilih disusun ulang secara acak kemudian digunakan untuk melakukan pengukuran di lapangan.

4. Menentukan sistem penskoran

Sistem penskoran yang digunakan instrumen tergantung pada skala pengukuran yang dipilih. Untuk skala Linkert, skor tertinggi tiap butir adalah 4 dan terendah 1. Apabila skala Thurstone, skor tertinggi tiap butir 7 dan terendah adalah 1. Adapun skala beda semantik skor tertinggi 7 dan skor terendah 1. Penjelasan lebih lanjut dapat dilihat pada BAB II.

5. Menelaah instrumen

Kegiatan menelaah instrumen adalah meneliti tentang beberapa aspek berikut.

- a. Apakah butir pertanyaan atau pernyataan sesuai dengan indikator.
- b. Apakah bahasa yang digunakan sudah komunikatif dan menggunakan tata bahasa yang benar.
- c. Apakah butir pertanyaan atau pernyataan tidak bias.
- d. Apakah format instrumen menarik untuk dibaca.
- e. Apakah jumlah butir sudah tepat sehingga tidak menjemukan dalam menjawabnya.

Telaah dilakukan oleh ahli. Hasil telaah dikonversikan dalam bentuk skor berdasarkan panduan penskoran. Untuk selanjutnya, hasil skoring digunakan sebagai acuan dalam menilai apakah instrumen yang terusun sangat baik, baik, cukup baik, tidak baik, atau sangat tidak baik. Berikut pedoman penskoran yang digunakan.

Tabel 3.1 Pedoman penskoran penilaian instrumen penilaian

Rentang skor	Nilai	Kategori
$x > (\bar{x} + 1,8 SBi)$	A	sangat baik
$(\bar{x} + 1,8 SBi) < X \leq (\bar{x} + 0,6 SBi)$	B	baik
$(\bar{x} + 0,6 SBi) < X \leq (\bar{x} - 0,6 SBi)$	C	cukup baik
$(\bar{x} - 0,6 SBi) < X \leq 2,6(\bar{x} - 1,8 SBi)$	D	kurang baik
$X \leq (\bar{x} - 1,8 SBi)$	E	sangat kurang baik

Keterangan:

\bar{x} = rerarata skor ideal ($\frac{1}{2}$ (skor maksimal + skor minimal))

SBi = simpangan baku ideal ($\frac{1}{6}$ (skor maksimal - skor minimal))

x = skor yang dicapai siswa

6. Melakukan ujicoba

Instrumen yang sudah ditelaah dan diperbaiki, selanjutnya digunakan untuk ujicoba. Ujicoba bertujuan untuk mengetahui karakteristik instrumen. Karakteristik yang penting adalah daya beda instrumen dan tingkat keandalannya. Semakin besar variasi jawaban tiap butir maka akan semakin baik instrumen.

7. Menganalisis instrumen

Setelah instrumen diujicobakan, maka langkah selanjutnya yaitu menganalisis instrumen. Hal ini dilakukan untuk mengetahui validitas dan reliabilitas butir angket.

Suatu pengukuran dikatakan memenuhi unsur validitas jika instrumen mampu mengukur apa yang seharusnya diukur. Artinya, jika instrumen tidak valid maka dikatakan instrumen

tidak mampu mengukur apa yang seharusnya diukur. Secara konseptual, ada 2 macam validitas, yaitu validitas isi (*content validity*) dan validitas konstruk (*construct validity*).

Validitas isi merupakan validitas yang berfungsi untuk memastikan bahwa ukuran telah cukup memasukkan sejumlah item yang representatif dalam menyusun sebuah indikator. Validitas ini terpenuhi oleh hasil penilaian ahli atau pakar pada bidang yang dituju.

Validitas konstruk membuktikan seberapa bagus hasil yang diperoleh dari penggunaan ukuran berdasarkan teori. Hal ini dinilai dengan *convergen validity* (instrumen yang memiliki korelasi tinggi) dan *discriminant validity* (variabel yang tidak berkorelasi).

Reliabilitas menunjukkan konsistensi dan stabilitas suatu skor. Reliabilitas mencakup 2 hal utama, yaitu stabilitas ukuran dan stabilitas internal ukuran. Stabilitas ukuran menunjukkan suatu ukuran untuk tetap stabil dan tidak rentan terhadap perubahan situasi apapun. Stabilitas internal ukuran merupakan indikasi homogenitas item-item yang ada dalam ukuran yang menyusun konstruk. Uji reliabilitas akan disudahi jika instrumen sudah memenuhi unsur reliabel, yaitu memiliki koefisien reliabilitas sebesar 0,7 (Mardapi, 2010).

8. Merakit instrumen

Kegiatan merakit instrumen merupakan tahap lanjutan setelah soal dinyatakan valid dan reliabel. Item yang belum memenuhi unsur valid maka akan dibuang atau diperbaiki. Item yang valid dan reliabel serta item hasil perbaikan kemudian disusun ulang untuk kemudian digunakan dalam melaksanakan pengukuran di lapangan.

9. Melaksanakan pengukuran

Pengukuran dilakukan dengan tujuan untuk mendapatkan data penelitian yang mengarah kepada pengambilan kesimpulan.

10. Menafsirkan hasil pengukuran

Nama lain dari menafsirkan hasil pengukuran adalah penilaian. Untuk melakukan penilaian maka diperlukan suatu kriteria berdasarkan skala dan jumlah butir yang digunakan.

Caranya yaitu dengan mencari skor keseluruhan yang dihasilkan dari instrumen yang telah diisi oleh responden. Untuk kemudian menyesuaikan dengan kriteria yang telah ditentukan. Berikut salah satu contoh penggunaan kriteria untuk mengukur sikap siswa.

Tabel 3.2 Kategorisasi sikap siswa

No	Skor siswa	Kategori sikap
1	$x \geq (\bar{x} + SBx)$	Sangat positif/sangat tinggi
2	$(\bar{x} + SBx) > x \geq \bar{x}$	Positif/tinggi
3	$\bar{x} > x \geq (\bar{x} - SBx)$	Negatif/rendah
4	$x < (\bar{x} - SBx)$	Sangat negatif/rendah

Keterangan:

\bar{x} = rerarata skor keseluruhan

SBx = simpangan baku

x = skor yang dicapai siswa

Demikian langkah pengembangan instrumen yang didasarkan pada sumber Mardapi (2010).

Retnawati (2013) memiliki pandangan yang berbeda dalam menentukan langkah-mengembangkan instrumen baik tes maupun nontes yang terbagi dalam 9 langkah, yaitu: (1) menentukan tujuan penyusunan instrument, (2) mencari teori yang relevan atau cakupan materi, (3) menyusun indikator butir instrument/soal, (4) menyusun butir instrument, (5) validasi isi, (6) revisi berdasarkan masukan

validator, (7) melakukan ujicoba kepada responden yang bersesuaian untuk memperoleh data respons peserta, (8) melakukan analisis (reliabilitas, tingkat kesulitan, dan daya pembeda), dan (9) merakit instrumen. Berikut penjelasan masing-masing langkahnya.

1. Menentukan tujuan penyusunan instrumen

Sebagai tahapan awal menyusun instrumen, perlu ditetapkan tujuan penyusunan instrumen. Tujuan penyusunan ini memandu teori untuk mengonstruksi instrumen, bentuk instrumen, penyekoran sekaligus pemaknaan hasil penyekoran pada instrumen yang akan dikembangkan. Tujuan penyusunan instrumen perlu disesuaikan dengan tujuan penelitian. Sebagai contoh, ketika peneliti akan mengetahui pengaruh pembelajaran roleplaying terhadap motivasi belajar dan kreativitas. Tentunya ada dua instrumen yang perlu dikembangkan, instrumen pengukur motivasi dan instrumen pengukur kreativitas.

2. Mencari teori yang relevan atau cakupan materi

Setelah tujuan penyusunan instrumen ditetapkan, selanjutnya perlu dicari teori atau cakupan materi yang relevan. Teori yang relevan digunakan untuk membuat konstruk, yaitu apa saja indikator suatu variabel yang akan diukur. Sebagai contoh motivasi belajar dan kreativitas mencakup beberapa indikator. Hal ini harus disebutkan.

3. Menyusun indikator butir instrumen/soal

Indikator soal ini ditentukan berdasarkan kajian teori yang relevan pada instrumen nontes. Indikator ini telah bersifat khusus, sehingga dengan menggunakan indikator dapat disusun menjadi butir instrumen. Biasanya aspek yang akan diukur dengan indikatornya disusun menjadi suatu tabel. Tabel tersebut kemudian disebut dengan kisi-kisi (*blue print*). Penyusunan kisi-kisi ini mempermudah peneliti menyusun butir soal.

4. Menyusun butir instrumen

Langkah selanjutnya adalah menyusun butir-butir instrumen. Penyusunan butir ini dilakukan dengan melihat indikator yang sudah disusun pada kisi-kisi. Pada penyusunan butir ini, peneliti perlu mempertimbangkan bentuknya. Misal untuk nontes akan

menggunakan angket, angket jenis yang mana, menggunakan berapa skala, penskorannya dan analisisnya. Pada penyusunan butir ini, peneliti perlu mempertimbangkan penskoran untuk tiap butir, sehingga memudahkan analisis.

5. Validasi isi

Setelah butir-butir soal tersusun, langkah selanjutnya adalah validasi. Validasi ini dilakukan dengan menyampaikan kisi-kisi, butir instrumen, dan lembar diberikan kepada ahli untuk ditelaah secara kuantitatif dan kualitatif. Tugas ahli adalah melihat kesuaian indikator dengan tujuan pengembangan instrumen, kesesuaian indikator dengan cakupan materi atau kesesuaian teori, melihat kesuaian instrumen dengan indikator butir, melihat kebenaran konsep butir soal, melihat kebenaran isi, bahasa dan budaya. Proses ini disebut dengan validasi isi dengan mempertimbangkan penilaian ahli (*expert judgement*). Jika validasi isi akan dikuantifikasi, peneliti dapat meminta ahli mengisi lembar penilaian validasi. Paling tidak, ada 3 ahli yang dilibatkan untuk proses validasi instrumen penelitian. Berdasarkan isian 3 ahli, selanjutnya penelitian menghitung indeks kesepakatan ahli atau kesepakatan validator dengan menggunakan indeks Aiken atau indeks Gregory.

6. Revisi berdasarkan masukan validator

Biasanya validator memberikan masukan. Masukan-masukan ini kemudian digunakan peneliti untuk merevisinya. Jika perlu, peneliti perlu mengkonsultasikan lagi hasil perbaikan tersebut, sehingga diperoleh instrumen yang benar-benar valid.

7. Melakukan ujicoba kepada responden yang bersesuaian untuk memperoleh data respons peserta

Setelah revisi, butir-butir instrumen kemudian disusun lengkap (dirakit) dan siap diujicobakan. Ujicoba ini dilakukan dalam rangka memperoleh bukti empiris. Ujicoba ini dilakukan kepada responden yang bersesuaian dengan subjek penelitian. Peneliti dapat pula menggunakan anggota populasi yang tidak menjadi anggota sampel.

8. Melakukan analisis instrumen (reliabilitas)
Setelah melakukan ujicoba, peneliti memperoleh data respons peserta ujicoba. Dengan menggunakan respons peserta, peneliti kemudian melakukan penskoran tiap butir. Selanjutnya hasil penskoran ini digunakan untuk melakukan analisis reliabilitas. Analisis reliabilitas dapat dilakukan dengan berbagai teknik. Analisis ini akan dibahas secara lebih jelas pada V.
9. Merakit instrumen
Setelah karakteristik butir diketahui, peneliti dapat merakit ulang perangkat instrumen. Pemilihan butir-butir dalam merakit perangkat ini mempertimbangkan karakteristik tertentu yang dikehendaki peneliti berdasarkan analisis instrumen. Setelah itu, diberi instruksi pengerjaan, peneliti kemudian dapat mempergunakan instrumen tersebut untuk mengumpulkan data penelitian.

Untuk lebih jelasnya, berikut contoh konkrit pengembangan instrumen angket.

B. Contoh Pengembangan Instrumen

Contoh 3.1 Pengembangan instrumen angket berdasarkan acuan sumber Mardapi (2008)

Pengembangan instrumen penilaian tanggung jawab dalam kegiatan pembelajaran (Sarah, Prasetyo, Wilujeng, 2018)

1. Menentukan spesifikasi instrumen
Variabel yang akan dicapai berupa penyusunan instrumen penilaian tanggung jawab dalam kegiatan pembelajaran. Instrumen yang digunakan adalah angket yang akan diisi oleh orang tua siswa. Hal ini dikarenakan tanggung jawab merupakan salah satu bentuk penilaian sikap sehingga membutuhkan proses panjang untuk mengukurnya. Selain itu, sikap tanggung jawab dipengaruhi banyak aspek yang menentukan, dari mulai keluarga, sekolah, hingga masyarakat.

2. Menulis instrumen

Instrumen yang disusun mengacu pada indikator variabel yang akan dikembangkan, yaitu tanggung jawab. Hal ini dapat dijabarkan dari teori dan hasil penelitian orang lain terkait variabel tanggung jawab.

Lickona (2013: 61) menyatakan bahwa tanggung jawab adalah perluasan dari sikap hormat. Jika kita menghormati orang lain, berarti kita menghargainya. Jika kita menghargai mereka, berarti kita merasakan tanggung jawab tertentu terhadap kesejahteraan mereka. Secara harfiah tanggung jawab berarti “kemampuan untuk menanggung”. Artinya, kita berorientasi pada orang lain, memberi perhatian kepada mereka, dan tanggap terhadap kebutuhan mereka. Tanggung jawab menekankan kewajiban-kewajiban positif untuk saling peduli terhadap satu sama lain.

Tillman (2004: 216) dan Salahudin (2013) memberi pengertian bahwa tanggung jawab adalah menerima kebutuhan dan melakukan tugas dengan sebaik-baiknya. Bertanggung jawab juga berarti melakukan kewajiban dengan sepenuh hati. Orang yang bertanggung jawab memiliki sesuatu yang bernilai untuk diberikan dan mengetahui cara berlaku adil. Pada hak terdapat tanggung jawab. Tanggung jawab bukan hanya suatu kewajiban, tetapi juga sesuatu yang membantu diri sendiri mencapai tujuan. Setiap orang dapat mengamati dunianya sendiri dan melihat keseimbangan antara hak dan kewajibannya. Tanggung jawab adalah menggunakan seluruh daya untuk perubahan yang positif. Jadi, inti tanggung jawab adalah keseimbangan antara hak dan kewajiban.

Munir (2010), Samani dan Hariyanto (2016) memaknai tanggung jawab sebagai kemampuan seseorang menjalankan kewajiban karena dorongan dalam dirinya, atau panggilan jiwa. Jadi, seseorang mengerjakan sesuatu bukan karena itu sebuah aturan, namun karena panggilan jiwa. Orang yang bertanggung jawab

merasa ada yang sesuatu yang kurang apabila tidak mengerjakan. Jadi, tanggung jawab adalah melakukan tugas sepenuh hati, bekerja dengan etos kerja tinggi, dan berusaha keras untuk mencapai prestasi.

Mu'in (2013) merangkum beberapa istilah dalam tanggung jawab.

- a. Tugas, yaitu apa yang harus dilaksanakan sebagai ganti dari apa yang telah didapatkan.
- b. Pembagian kerja, yaitu tanggung jawab untuk melaksanakan kerja sesuai rencana yang telah ditetapkan.
- c. *Relationship obligation* (kewajiban dalam hubungan), bahwa tanggung jawab penting dalam suatu hubungan. Contohnya, kerja kelompok memerlukan tanggung jawab dan kerja sama dalam menyelesaikan tugas.
- d. Dapat dipertanggungjawabkan, yaitu keadaan yang bisa dimintai tanggung jawab dan bisa dipertanggung jawabkan. Misalnya, dapat mempertanggungjawabkan perkataan yang telah diucapkan.
- e. Ketekunan dan sifat rajin: orang yang rajin dan tekun biasanya orang yang bertanggung jawab.
- f. *Teamwork* (kerja kelompok); orang yang dapat bekerja sama dalam menyelesaikan masalah kelompok adalah orang yang bertanggung jawab.
- g. *Accountability* (dapat dipertanggungjawabkan), keadaan yang bisa dimintai tanggung jawab dan bisa dipertanggungjawabkan.
- h. *Self-motivated* (motivasi diri): orang yang bertanggung jawab itu memiliki kemampuan motivasi diri sendiri.

Berdasarkan berbagai penjelasan, maka disimpulkan bahwa tanggung jawab adalah sikap dan perilaku seseorang untuk melaksanakan tugas dan kewajibannya sebagaimana yang seharusnya dilakukan terhadap diri sendiri, masyarakat, dan lingkungan (alam, sosial, dan budaya), negara, dan Tuhan.

Indikator tanggung jawab meliputi: (1) mengerjakan tugas semestinya dan (2) menerima resiko dari tindakan yang dilakukan. Masing-masing indikator kemudian dijabarkan pada instrumen angket dengan kisi-kisi terdapat pada Tabel 1 yang kemudian dijabarkan dalam bentuk instrumen angket (Tabel 2).

Berikut 3.3 kisi-kisi sikap tanggung jawab

No	Indikator	No butir		Jmh
		+	-	
1	Mengerjakan tugas semestinya	3,9,5	4,8,10	6
2	Menerima resiko dari tindakan yang dilakukan	2,6	1,7	4

Tabel. 3.4 Angket tanggung jawab

No	Pernyataan	Skor			
		1	2	3	4
1	Anak saya bicara berdasarkan fakta di lapangan				
2	Anak saya tidak meminta maaf atas kesalahan yang mengakibatkan kelompoknya gagal menyelesaikan tugas				
3	Anak saya mengakui kesalahan hingga mengakibatkan kelompoknya gagal dalam menyelesaikan tugas				
4	Anak saya rajin mengerjakan tugas sekolah secara individu maupun berkelompok				
5	Anak saya mengabaikan apa yang ditugaskan guru				
6	Anak saya mendiskusikan sesuatu yang belum diketahuinya dengan siapa saja yang dianggap dapat membantunya				
7	Anak saya cepat bertindak ketika diminta bantuan				

No	Pernyataan	Skor			
		1	2	3	4
8	Anak saya dapat menjelaskan masalah yang dialaminya dengan baik				
9	Anak saya cepat dalam menyelesaikan tugas				
10	Anak saya mengerjakan apa yang menjadi tugasnya				
	Skor total				

3. Menentukan skala pengukuran
Instrumen penilaian tanggung jawab menggunakan angket skala Likert dengan 4 option, yaitu 1 = tidak pernah, 2 = jarang, 3 = sering, dan 4 = selalu.
4. Menentukan sistem penskoran
Berdasarkan skala Likert pada instrumen dan jumlah item angket, maka skala maksimum yang dapat diperoleh responden sebanyak 40, sebaliknya skala minimum sebanyak 10.

Kegiatan ke-5 yaitu menelaah instrumen hingga ke 10 yaitu menggunakan instrument akan diuraikan di BAB berikutnya.

Contoh 3.2 Pengembangan instrument angket berdasarkan acuan sumber Mardapi (2013)

Contoh 2 diadopsi dari Penelitian Sarah, Prasetyo, dan Wilujeng (2019) berjudul “*Development of Living Values (Honesty) Assessment Instrument in Learning Process*”. Instrumen ini membahas tentang Pengembangan instrumen living values (kejujuran) bagi siswa Sekolah Menengah Atas (SMA). Berikut langkahnya.

1. Menentukan tujuan penyusunan instrumen

Tujuan penyusunan instrumen ini adalah mengembangkan instrumen *living values* (kejujuran) dalam proses pembelajaran di SMA. Kejujuran merupakan salah satu nilai yang ada dalam

living values. Kejujuran merupakan salah satu bagian dari penilaian afektif. Oleh karena itu, angket menjadi salah satu instrumen yang dikembangkan untuk menilai kejujuran seseorang.

2. *Mencari teori yang relevan atau cakupan materi,*

Karena yang dikembangkan adalah angket kejujuran dalam proses pembelajaran di SMA, maka cakupan materi diambil dari teori yang relevan dari buku referensi, jurnal-jurnal ataupun penelitian yang telah terdahulu.

Berikut alur pengembangan angket kejujuran berdasarkan teori dan referensi yang digunakan dalam mengembangkan instrumen *living values* (kejujuran) (diambil dari Sarah, Prasetyo, & Wilujeng, 2019).

Living values adalah dasar dari pembentukan karakter. Nilai ini mudah diinternalisasi dan diterapkan. Menurut Tillman (2004: 10), *living values* merupakan program pendidikan nilai yang komprehensif. Program ini menawarkan pelatihan, metodologi praktis dan berbagai kegiatan nilai pengalaman kepada pendidik, fasilitator, orang tua, dan pembimbing untuk membantu mereka memberikan kesempatan bagi anak-anak guna menggali dan mengembangkan nilai-nilai universal yang merupakan nilai-nilai kunci pribadi dan sosial.

Penanaman *living values* dapat dilakukan melalui pendekatan komprehensif dari sisi metode meliputi: (1) inkulkasi, (2) keteladanan, (3) fasilitasi, dan (4) pengembangan keterampilan. Inkulkasi berupa penanaman nilai. Hal ini dapat dilaksanakan melalui strategi penggunaan fiksi dan nonfiksi, hadiah dan apresiasi, simbol-slogan-poster, pengajaran empati, kegiatan ekstrakurikuler, dan pengembangan *self-esteem* (kesadaran akan harga diri). Keteladanan dilakukan dengan cara menunjukkan kepada peserta didik cara terbaik dalam mengatasi masalah. Berbagai strategi yang dapat dilakukan untuk memberikan keteladanan antara lain: berbagi perasaan, berbagi pengalaman, nara sumber, dan menghindari hipokrasi

(kemunafikan). Fasilitasi melatih peserta didik mengatasi masalah. Strategi pembelajaran yang dapat dilakukan melalui fasilitasi antara lain: menentukan prioritas, wawancara, puisi sarat nilai, diskusi dilema moral, evaluasi diri, dan debat-debat tentang isu-isu kontroversial. Adapun strategi yang dikembangkan untuk menghidupkan nilai-nilai melalui keterampilan antara lain: berpikir kritis, berpikir kreatif, berkomunikasi secara jelas, menyimak, bertindak asertif, dan menemukan resolusi konflik yang secara ringkas disebut keterampilan akademik dan keterampilan sosial. (Kirschenbaum, 1995).

Tujuan pendidikan berbasis living values yaitu: (1) membantu individu memikirkan dan merefleksikan nilai-nilai dan mengekspresikan implikasi praktisnya dalam kaitannya dengan pribadi, orang lain, masyarakat, dan dunia pada umumnya; (2) memperdalam pemahaman, motivasi, dan tanggung jawab yang berkaitan dengan membuat pilihan-pilihan pribadi dan sosial yang positif; (3) menginspirasi individu dalam memilih nilai-nilai pribadi, sosial, moral, dan spiritual serta menyadari metode praktis untuk mengembangkan dan memperdalam; (4) mendorong pendidik melihat pendidikan sebagai sarana memberikan filosofi hidup untuk memfasilitasi pertumbuhan peserta didik secara keseluruhan, mengembangkan, dan memberikan pilihan sehingga dapat mengintegrasikan diri pada masyarakat dengan rasa hormat, percaya, dan sarat akan tujuan. (Tillman (2012: 8-9)

Ada 13 nilai hidup yang dapat dikembangkan di Indonesia (Komalasari, 2012). Salah satunya adalah kejujuran. Kejujuran menurut Tillman (2004: 120) yaitu mengatakan kebenaran. Saat seseorang jujur, maka orang tersebut akan berpikiran jernih. Kejujuran juga berarti tidak kontradiksi dalam pikiran, kata, atau tindakan. Pikiran, kata-kata, tindakan yang jujur akan menciptakan harmoni. Kejujuran adalah kesadaran akan apa yang benar dan sesuai dengan perannya, tindakannya, dan hubungannya. Dengan kejujuran, tidak ada kemunafikan atau kepalsuan yang menciptakan kebingungan dan ketidakpercayaan

dalam pikiran dan hidup orang lain. Kejujuran membuat integritas dalam hidup, karena apa yang ada di dalam dan di luar diri adalah cermin jiwa. Kejujuran digunakan pada apa yang dipercayai seseorang. Ada hubungan yang erat antara kejujuran dan persahabatan. Ketamakan kadang kala bersumber dari ketidakjujuran seperti dilakukan koruptor. Orang yang jujur mengetahui bahwa kita semua saling berhubungan. Jujur pada diri dan dalam menghadapi tugas, akan mendapatkan kepercayaan dari dan mengilhami orang lain.

Jujur berarti bersikap terbuka, dapat dipercaya, dan mengatakan kebenaran. Bila orang jujur, mereka bisa diandalkan untuk tidak berbohong, menipu, atau mencuri, dan melakukan tindak korupsi. Kejujuran adalah mengatakan yang sebenarnya. Seseorang yang jujur akan mengakui kesalahan bahkan ketika orang lain mengetahui mungkin akan marah atau kecewa. Bersikap jujur berarti tidak berpura-pura untuk menjadi sesuatu yang tidak diinginkan. Dengan kejujuran, seseorang bisa mempercayai hal apa adanya (*The random acts of kindness foundation*, 2013). Komalasari (2012: 249) dan Koesoema (2015) mendefinisikan kejujuran sebagai sebuah perilaku yang dilakukan untuk menjadikan seseorang sebagai orang yang terpercaya dengan kata-kata, tindakan, dan pekerjaan baik terhadap diri sendiri maupun orang lain.

Hasil penelitian Mazar, Amir, dan Ariely, 2008: 643) menunjukkan bahwa orang berperilaku tidak jujur cukup untuk mendapatkan keuntungan tapi cukup jujur untuk menipu diri mereka sendiri atas integritas mereka sendiri. Dalam kehidupan sehari-hari, tidak ada keraguan bahwa ketidakjujuran lazim (sudah menjadi hal yang umum). Oleh karena itu, diperlukan tindakan yang efektif dan efisiensi untuk mencegah ketidakjujuran. Salah satunya dengan memahami mekanisme yang berbeda yang mendasari perilaku dalam situasi tertentu.

Perilaku jujur dapat dihidupkan melalui proses pembelajaran. Tillman (2004: 131-132) memberi contoh desain pembelajaran sains untuk menghidupkan kejujuran melalui beberapa

pertanyaan. Contoh sederhana yaitu pertanyaan mengenai apa akibat yang terjadi saat seorang ilmuwan mencuri ide ilmuwan lain. Pertanyaan lainnya seperti apa saja etika dalam sains; apa yang akan terjadi jika data hasil penelitian salah; bagaimana hubungan antara ketidakjujuran, integritas dan menghargai diri. Beberapa pertanyaan di atas mengarahkan peserta didik untuk memiliki sikap jujur melalui pembelajaran fisika.

Sikap jujur akan dapat hidup lama dalam diri peserta didik jika dilandasi dengan ketakwaan terhadap Tuhan Yang Maha Esa. Ismail (2009: v) menyatakan bahwa takwa merupakan nilai atau akumulasi dari nilai-nilai islam. Madjid (1992) menyatakan takwa membentuk kualitas pribadi orang yang beriman. Adapun Shihab (2007) menyatakan bahwa takwa melindungi seseorang dari siksa dan bencana.

Ismail (2009: xi-xvi) menyatakan bahwa takwa berakar pada empat pilar. (1) Kesadaran ketuhanan. Artinya, kesadaran pada seseorang bahwa Tuhan selalu hadir menyertai hidupnya serta mengawasi semua tingkah lakunya, sehingga yang bersangkutan takut dan malu berbuat dosa. (2) Semangat ibadah dan ketaatan yang tinggi kepada Tuhan. Artinya, orang yang memiliki kesadaran ketuhanan yang tinggi dengan sendirinya memiliki semangat ibadah dan tingkat kepatuhan yang tinggi kepada Tuhan. Kepatuhan menyangkut 2 aspek, yaitu lahiriyah dan batiniyah. Lahiriyah berupa kepatuhan kepada hukum-hukum Tuhan dan batiniyah berupa kesucian niat dan sikap yang tulus dalam beragama. (3) Semangat kemanusiaan dan kesalehan sosial. Maksudnya, agama diturunkan bukan untuk Alloh, akan tetapi untuk manusia. Jadi, manusia harus beragama dan melaksanakan perintah-perintah agama, bukan untuk Tuhan namun untuk kebaikan dan kemaslahatan manusia itu sendiri dalam kehidupannya di dunia dan akhirat. (4) Kualitas moral dan keluhuran budi pekerti. Artinya, orang yang takwa adalah manusia dengan kualitas moral yang tinggi.

Berdasarkan berbagai uraian sebagaimana, maka yang dimaksud dengan kejujuran adalah sikap dan perilaku seseorang

yang didasarkan pada upaya menjadikan dirinya selalu dapat dipercaya dalam perkataan dan perbuatan. Indikator kejujuran yaitu mengatakan sesuatu apa adanya (tidak berbohong), tidak curang, mampu menjaga kepercayaan orang lain, dan menuliskan data percobaan sesuai hasil yang diperoleh.

Karakter adalah bagian dari domain afektif (Mardapi, 2017). Artinya, nilai-nilai hidup juga merupakan bagian dari aspek afektif. Domain afektif mencakup 5 tahap, yaitu menerima, merespons, menilai, organisasi, dan karakterisasi. Menerima berarti siswa memiliki keinginan. Menanggapi berarti bahwa siswa tidak hanya memiliki keinginan tetapi juga bereaksi. Menilai berarti bahwa siswa memiliki perilaku yang konsisten dan stabil sebagai hasil dari pembelajaran.

Hasil pembelajaran pada tahap organisasi adalah dalam bentuk konseptualisasi nilai. Dalam bidang karakterisasi, siswa memiliki sistem nilai yang mengontrol perilaku sampai waktu tertentu untuk membentuk gaya hidup (Krathwohl et al., 1964). Metode pengukuran domain afektif termasuk observasi dan pelaporan diri (Andersen, 1981). Penilaian dalam domain afektif membutuhkan data kuantitatif atau kualitatif. Data kuantitatif diperoleh melalui pengukuran. Hasilnya dalam bentuk angka. Data kualitatif diperoleh melalui instrumen non-tes, seperti pedoman observasi dan kuesioner. Berdasarkan studi teoritis, penelitian ini mengembangkan 3 instrumen, yaitu lembar observasi, lembar kuesioner, dan lembar pedoman wawancara. Lembar observasi mengukur kejujuran siswa dalam proses pembelajaran. Lembar kuesioner diisi oleh orang tua siswa. Pedoman wawancara berisi serangkaian pertanyaan tentang kejujuran siswa. Wawancara dilakukan dengan siswa. Pada penelitian ini, hanya akan membahas tentang instrumen angket.

Tabel 3.5 Indikator dan instrumen penilaian kejujuran

No	Indikator	Insrumen
1	Mengatakan sesuatu apa adanya (tidak berbohong)	Lembar angket Lembar observasi
2	Tidak curang	Lembar angket Lembar observasi Lembar wawancara
3	Mampu menjaga kepercayaan orang lain	Lembar angket Lembar observasi
4	Menuliskan data percobaan sesungguhnya	Lembar observasi

3. Menyusun indikator butir instrumen/soal

Berdasarkan teori-teori yang relevan, dikonstruksikan indikator-indikator kejujuran. Untuk memperjelas tiap indikator, berikut penyusunan indikator butir instrumen.

Untuk memperoleh hasil penelitian yang terpercaya, angket penilaian kejujuran disusun dalam bentuk pernyataan positif dan negatif dan diletakkan secara acak untuk mendapatkan jawaban responden yang jelas dan bukan menebak. Pernyataan disusun menggunakan 4 skala Likert, yaitu 1 = tidak pernah, 2 = jarang, 3 = sering, 4 = selalu.

Lembar observasi kejujuran tersusun dari 3 item dengan skala 1 – 4 berdasarkan rubrik. Semakin banyak aktivitas yang mengarah ke pencapaian kejujuran, skor responden semakin mendekati 4. Sebagian indikator lembar observasi dikembangkan menjadi subindikator.

Tabel 3.6 Indikator item angket kejujuran

No	Indikator	Indikator pernyataan	Nomor item
1	Mengatakan sesuatu apa adanya (tidak berbohong)	Berbicara berdasarkan fakta di lapangan	1
		Tidak berbicara berdasarkan fakta	4
2	Tidak curang	Orang lain mengerjakan tugas	2
		Mengerjakan tugas sendiri	5
3	Mampu menjaga kepercayaan orang lain	Mengembalikan barang orang lain dalam keadaan baik dan bersih	6
		Tidak mengembalikan barang orang lain	3
		Meminta izin sebelum menggunakan barang orang lain	7
		Tidak meminta izin saat menggunakan barang orang lain	8

4. Menyusun butir instrumen

Berdasarkan kisi-kisi yang ada, langkah selanjutnya yaitu menyusun butir instrumen.

Berikut butir angket untuk mengukur kejujuran siswa SMA melalui kegiatan pembelajaran.

Identitas Responden

Nama Orang Tua :

Nama Siswa :

LEMBAR ANGGKET LIVING VALUES (KEJUJURAN)

Petunjuk

1. Pernyataan-pernyataan berikut ini berkaitan dengan kebiasaan putra/putri Bapak/Ibu sehari-hari. Isilah pilihan pada lembar jawaban yang tersedia dengan memberi tanda cek (√) pada kolom yang sesuai.
2. Isilah setiap nomor dengan kenyataan karena hal ini sama sekali tidak berpengaruh terhadap penilaian putra/putri Bapak/Ibu di sekolah. Atas partisipasinya kami ucapkan terima kasih.

Keterangan

1 = tidak pernah

3 = sering

2 = jarang

4 = selalu

No	Pernyataan	Skor			
		1	2	3	4
1	Anak saya bicara berdasarkan fakta di lapangan				
2	Anak saya menyuruh orang lain untuk mengerjakan tugasnya				
3	Anak saya tidak mengembalikan barang orang lain yang dipinjamnya				
4	Apa yang dibicarakan anak saya berbeda dengan fakta				
5	Anak saya mengerjakan tugasnya sendiri				
6	Anak saya mengembalikan barang orang lain yang dipinjam dalam keadaan baik dan bersih				
7	Anak saya minta izin ketika mau memakai barang orang lain				
8	Anak saya tidak meminta izin terlebih dahulu saat mau memakai barang orang lain				
	Skor total				

Penerapan Isu Lingkungan 3.1

Penelitian membahas tentang pengembangan instrument *living values* (kejujuran) dalam proses pembelajaran

Sumber: Sarah, S. Prasteyo, Z. K., & Wilujeng, I. (2019). Development of Living Values (Honesty) Assessment Instrument in Learning Process. *International Journal of Instruction*, 12(4), 443-458. <https://doi.org/10.29333/iji.2019.12429a>

Abstract

The cheating among students shows a low value of honesty. A learning process in high school loaded with living values, namely honesty. This study aims to produce instruments to measure honesty in the learning process that is valid and reliable. The research used for research and development methods with 8 steps, namely (1) determining instrument specifications, (2) compiling, (3) choosing a measurement scale, (4) scoring, (5) examining, (6) limited experiment, (7) analyzing, and (8) concluding. The first trial involved 1167 parents of high school students in x class. A second trial involved 1261 parents of high school students on xi and xii class. Respondents lived in Wonosobo Regency, Indonesia. The research instruments included a questionnaire, observation, and interview instruction sheets. Data were analysed by expert judgment and explanatory factor analysis. Research produces valid and reliable instruments to measure honesty. A questionnaire with 4 items have Alpha's Cronbach coefficient 0.840. The observation consisted of 3 items with very good criteria. The interview guide has fulfilled the content validity after being corrected based on expert advice. The implication of the study is the students' honesty can be measure early so as not to develop into dishonest behaviour and lead to criminal acts.

Keywords: honesty, assessment, instrument, learning process, *living values*.

INTRODUCTION

The people of Indonesia highly uphold the noble values of the nation such as honesty, politeness, togetherness, religion, and others which are all listed in the state foundation, namely Pancasila. The noble values have taken root in the attitudes and behaviors of life forming character (Zuchdi, 2012). However, along with the development of the era the character is increasingly fading. Indonesian people hope education can create graduates that benefit the community. It can be fulfilled if graduates have good character (Mardapi, 2017). Therefore, starting in 2013 until now the Indonesian National Education system seeks to make Indonesian people not only have intelligence, but also good spiritual, personal, and self-control characters (Hidayati, et al., 2018). This step is realized by inclusion the achievement of characters in the curriculum (Regulation of the Minister of National Education-Kemendiknas, 2010).

The character is a moral quality and is a consideration for someone in making decisions, acting, and behaving that are relatively fixed (Mardapi, 2017). The character is not a variable that can be achieved quickly. It takes a long time to reach it. Characters are formed by values, autonomy, and determination and loyalty (Zuchdi et al., 2013). Pancasila, the basis of the Indonesian state explains the character of the Indonesian nation which turns out to be in line with the achievement of living values (Komalasari & Saripudin, 2018). Living values are the basic values of character building. It is easy to be internalized and implemented. Komalasari (2012) & Wening (2012) said that living values contributed 26% to the formation of national character. Living values through habitual activities in school can improve student character (Komalasari et al., 2014). Tillman (2004) fosters living values by encouraging students to see their actions in others and how they can make a difference.

Living values are closely related to culture. Culture is used in various contexts which means many different things and generally refers to habit, practices, language, values that grow in society (Daskon, 2015). Komalasari (2012) shows 13 living values that can be developed in Indonesia namely religious, honesty, tolerance, ethical

behavior, discipline, working hard, creative, independent, democratic, loving the country, respecting achievement, working together, and responsible. Among the 13 living values, honesty is the most urgent living values for all elements of the nation.

Dishonesty practices such as plagiarism and cheating behavior are prevalent in society (Mazar et al., 2008). Lickona (2013) states that one sign of the nation's destruction is to cultivate dishonesty. Honesty is not the only attitude, but dignity, self-esteem, one's identity and national identity (Muhasim, 2017).

Plagiarism practices occur in almost all countries in the world such as America, Bangladesh, India, Pakistan, Malaysia, Africa, Norway, Kenya, Iran, and America (Moten, 2014). Bunn, Caudill, and Gropper (1992) shows that the majority of students admits to cheating/plagiarism while taking college. Cheating among students was no less alarming. The results of the Josephson Institute Center for Youth survey of 23,000 public and private high school students found that 51% of students cheated during the year-end test. The survey also found that 74% of students copied the homework of other students (Ehrlich & Fu, 2013).

Based on the information, dishonesty has existed from ancient times until now. Rules for perpetrators of dishonesty already exist. The punishment for the perpetrator of dishonesty has been carried out. However, the practice of dishonesty still continues. Honesty is an attitude or behavior that will develop into a habit. It takes a long time (Shalvi et al., 2012). The practices of dishonesty start from simple attitudes and behavior. It should be detected as early as possible so as not to develop. The lie detector (polygraph) is only used to detect the lies of perpetrators of crime in the legal sphere. Novodvorsky (1993) developed an instrument for measuring attitudes (honesty) in science at the high school level using questionnaires filled with students. The limited detection of dishonesty is feared to be less able to reveal someone's honesty appropriately. Therefore, it is necessary to assess honesty using a valid assessment instrument (Mardapi, 2017). LaForgia (1988) states that "researchers need to provide reliable and valid and unsustainable evaluation tools as they did in the past by

developing instruments that will hardly meet simple criteria". This study aims to develop instruments to measure honesty through the learning process. It takes learning processes from the side of physics learning because it is at the core. The process includes observation, demonstration, and experiment (Klieger & Sherman, 2015; Ogunleye & Babajide, 2011). When making observations, careful and honest behavior is needed to produce significant data (Scheider, 1980). The honesty assessment instrument is expected to be a reference for everyone, especially teachers, to check the honesty of their students in the school environment and people in the family environment. Thus, measurement of honesty will get complete information.

METHOD

Type of Research

This study is a research and development (R & D) model by Mardapi (2017) design in 10 steps. The study only takes eight stages because the process of character formation takes a long time (Mardapi, 2017; Zuchdi, et al., 2013). The eight steps were (1) determining instrument specifications, (2) compiling, (3) choosing a measurement scale, (4) making scoring, (5) examining, (6) doing a limited experiment, (7) analyzing, and (8) concluding.

Instrument specifications include setting goals and framework. The goals include data types and instruments. The framework was formed by constructing instruments based on conceptual definitions. It took from theories. Conceptual definitions developed into operational definitions and then developed into many indicators. It was used as guidelines to make items. The instrument equips with a scale of measurement and score. It was reviewed by experts to find out the quality of the items. After being valid, it was tested to the respondent to fulfill the content validity by a statistical test. The results used to analyze and produce a final instrument. Limited experiments are only conducted on questionnaires. This is not done on the observation sheet and interview guidelines because it is constrained by bureaucracy and takes a long time.

Respondents

Wonosobo is regency in Central Java, Indonesia with a high percentage of poor people (20.32) and poverty depth index (1.10) in 2013-2017. On education, the net enrollment rate of high school students is low. This criterion comes from a comparison of data from other regencies in Central Java (BPS-Statistics of Wonosobo Regency, 2018).

There are sixteen high schools in Wonosobo. They are eleven public high schools and five private high schools. In this research, the school was chosen randomly by drawing. They are seven public high schools and two private high schools. On the first trial, the respondents were the parents of students who learn physics in the school of x class. On the second trial, parents of a student who learn physics in the school of xi and xii class were chosen to be respondents. Physics learning has chosen because of the characteristics of the learning process.

Data collection instruments

Data collected through a questionnaire, observation sheets, and an interview. The questionnaire consists of eight items. It contains questions about students' honesty based on the parents' observations at home. The observation sheet measures students' honesty while doing the learning process. Observations were carried out by two observers. The interview sheet is semi-closed style. The interview was aimed at many students. Figure 1 clarifies the type of instrument in this study.

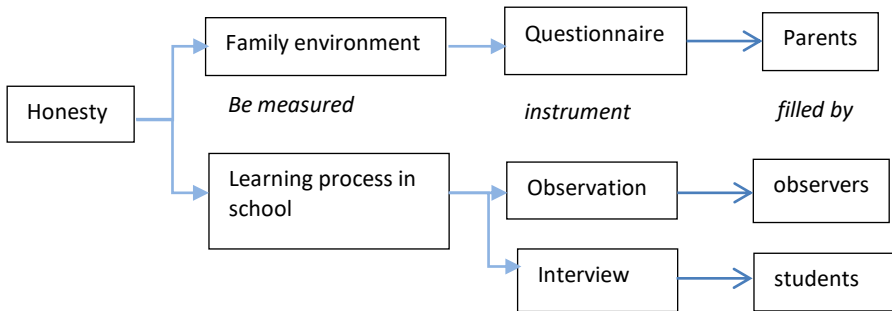


Figure 1
Diagram of how to measure honesty

Data Analysis

Data analyzed through judgment by six experts. They are three universities, academics and three physics teachers. The academics are physics learning, linguists, and measurement experts. The physics teachers came from public high school (two persons) and a private high school (one person). They assess the instrument, including content, language, and construction. The measurement results are converted to four scale values by reference to Table 1. Instruments fulfill the content validity by the expert judgment if the results of the assessment indicate the criteria of "very good" and "good"

Table 1
Conversion of measurement results

No	Interval score	Category
1	$x \geq \bar{x} + 1.5SBx$	Very good
2	$\bar{x} + 1.5SBx > x \geq \bar{x}$	Good
3	$\bar{x} > x \geq \bar{x} - 1.5SBx$	Not good
4	$x < \bar{x} - 1.5SBx$	Very bad

Description (Mardapi, 2008):

\bar{x} = mean ideal = $\frac{1}{2}$ (ideal highest score + ideal lowest score)

SBx = standard deviation ideal score (SB_i) = $\frac{1}{6}$ (ideal highest score - ideal lowest score)

x = score

Especially in the questionnaire data analysis was continued by explanatory factor analysis (EFA). EFA aims to find out the factors that construct honesty while proving the construct validity of the questionnaire. It was analyzed by SPSS version 25.00. Chi-square at the Barlette Test showed the adequacy of the sample. The sample was sufficient when the significance value is less than 0.01. This data are reinforced by KMO MSA when values above 0.50 (Hair, et al., 2010). Questionnaire reliability knows by Cronbach's Alpha coefficient.

FINDINGS

Determining instrument specifications (step 1)

Living values are the basis of character building. This value is easy to internalize and implement. There are 13 living values can be developed in Indonesia (Komalasari, 2012). One of them is honesty. Honesty is telling the truth. It means no contradiction in thoughts, words, or actions. Honesty makes integrity in life. There is a close relationship between honesty and friendship. Honest people know that everyone is interconnected. Be honest with yourself and in the face of assignments, will gain trust from and inspire others (Tillman, 2004). Komalasari (2012) defines more specifically. Honesty is a behavior that makes someone trusted by words, actions, and work. Mazar et al. (2008) said that dishonesty people are quite profitable but are cheating themselves.

Honesty is the attitude and behavior of a person to make himself always trustworthy in words and deeds. Indicators of honesty include: (1) not lying, (2) not cheating, (3) maintaining the trust of others, and (4) writing actual experimental data.

A character is part of the affective domain (Mardapi, 2017). That is, living values are also part of the affective aspect. Affective domain includes 5 stages, namely receiving, responding, valuing, organization, and characterization. Receiving means students have desires. Responding means that students not only have desires but also react. Valuing means that students have consistent and stable behavior as a result of learning. Learning outcomes at the organizational stage are in the form of value conceptualization. In the realm of

characterization, students have a value system that controls behavior until particular time to form a lifestyle (Krathwohl et al., 1964). Methods of measuring affective domains include observation and self-reporting (Andersen, 1981). Assessment in the affective domain requires quantitative or qualitative data. Quantitative data are obtained through measurements. The results are in the form of numbers. Qualitative data are obtained through non-test instruments, such as observation guidelines and questionnaires. Based on theoretical studies, this study developed 3 instruments, namely observation sheet, questionnaire sheet, and interview guideline sheet. Observation sheet measured students' honesty in the learning process. The questionnaire sheet is filled by student' parents. The interview guideline contains a series of questions about student honesty. Interviews are conducted with students.

Table 2

The framework of honesty instruments in the learning process

No	Indicators	Instruments
1	not lying	Questionnaire sheet
		Interview sheet
2	not cheating	Questionnaire sheet
		Interview sheet
		Observation sheet
3	maintaining the trust of others	Questionnaire sheet
		Interview sheet
4	writing experiment data	Observation sheet

Compiling (Step 2)

Instrument was making by the reference framework in Table 3. An honesty questionnaire consists of 8 items in the form of positive and negative statements and spread randomly, so that the respondent does not guess. The observation sheet and interview guideline sheet are each composed of 3 items.

Table 3
Items on the questionnaire

Indicators	Item indicators	Item number
Not lying	Speaking based on facts in the field	1
	No Speaking based on the fact	4
Not cheating	Has someone else done the work	2
	Working on tasks by their self	5
Maintaining the trust of other	Returning other people's object by properly and cleanly	6
	No returning other people's object	3
	Ask permission before using other people's items	7
	No permission when using other people's items	8

Table 4
Items on the observation sheet

Indicators	Item indicators	Item number
Not cheating	No copying data from other people or groups	1
Maintaining the trust of other	Working on problems based on their abilities	2
Writing data from experiment	Writing data according to experimental results	3

Table 5
Items on the interview guideline sheet

Indicators	Questions item	Item number
Not lying	Are there any dishonest behaviors that you have done? If yes, give an example.	1
Not cheating	Are there cheating behaviors that you have ever done? If yes, what kind of cheating behavior have you done?	2
Maintaining the trust of other	How do you maintain the trust of others?	3

Choosing a measurement scale (step 3) and scoring (step 4)

Questionnaire items are arranged using a Likert scale in the range of 1 to 4, namely 1 = never, 2 = rarely, 3 = often, 4 = always. Never are means never doing. Rarely means ever doing, but with a long time span, such as 1 or 2 times in 1 month. Often means doing something several times, which is 1 to 2 times a week. It always means doing things continuously. That is, the behavior has become a habit. The observation sheet is made using a rubric with 4 scales. Score 1 if there is no one indicator fulfilled. Score 2, if there is 1 indicator is fulfilled. Score 3, if there is 2 indicator is fulfilled. Score 4, if there is 3 indicator is fulfilled.

Examining (step 5)

The examining instrument aims to find out the quality of questions or statements based on indicators by expert judgment.

Table 6
Results of a questionnaire sheet review by expert judgment

Validators	Aspect	Average score	Maximum Score	Category
Universities academics	Content	4.33	5	Very good
	Language	24.67	25	Very good
	Construction	14.67	15	Very good
Physics teachers	Content	4.00	5	Very good
	Language	21.50	25	Very good
	Constructions	13.23	15	Very good

Table 6 shows that universities academics stated that the questionnaire was in the criteria of "very good" in content, language, and construction aspects. The physics teachers gave the same assessment with the university academics about the questionnaire. That is, the questionnaire sheet has fulfilled content validity.

Table 7
Results of an observation sheet review by expert judgment

Validators	Aspect	Average score	Maximum score	Category
Universities academics	Content	4.33	5	Very good
	Language	14	15	Very good
	Construction	13.33	15	Very good
Physics teacher	Content	4.5	5	Very good
	Language	13.5	15	Very good
	Constructions	15	15	Very good

Table 7 shows that universities academics and physics teacher consider the observation sheet to be in the "very good" category in content, language, and construction aspects. That is, the observation sheet has fulfilled content validity.

Table 8
Results of interview guideline validation by expert judgment

Validators	Aspect	Average score	Maximum score	Category
Universities academics	Content	2.67	5	Not good
	Language	15	15	Very good
	Construction	5	5	Very good
Physics teachers	Content	4.5	5	Very good
	Language	13.5	15	Very good
	Constructions	4.5	5	Very good

Table 8 shows that all experts gave an "very good" assessment of all aspects of the interview guide except the contents of universities academics by "not good" criteria. According to universities academics, the questions are not detailed enough to gather information about students' honesty. Therefore, questions are corrected so that more detailed information about student honesty.

Doing limited experiment (step 6)

The limited experiment was only carried out on questionnaires filled out by parents. Questionnaires were distributed to high school in the Wonosobo District by physics teachers' helping. A total of 9 high

schools were involved in this study. They are 7 public high schools and 2 private high schools. The questionnaire has spread for one week. After one week a questionnaire was taken back. There are 1240 questionnaires distributed, but 1184 were returned. After going through the sorting stage, 1167 questionnaires fulfilled the requirements for analysis. Sort was done because some questionnaires contain the same answer for all items. This indicates the respondent answered the questionnaire carelessly. Some respondents fill only a few items, but the others are left blank.

Analyzing instruments (step 7)

The adequacy of the sample shows the Chi-square value in the Bartlet test of 1013.420 with degrees of freedom 28 and p-values less than 0.01. That is, the sample size of 1167 used is sufficient. It is corroborated by the Kaiser Meyer Olkin measure of adequacy sampling (KMO) of 0.721 greater than 0.5.

The analysis was continued to find out the forming factors of honesty. The anti-image correlation (AIC) in Table 9 shows the variables that are feasible to use in carrying out factor analysis. It can be seen from the value of a Measure of Sampling Adequacy (MSA) by “a” code. MSA values above 0.50 indicate items are feasible to use factor analysis (Hair, 2010).

Tabel 9
Anti-image Matrices

	Item2	Item3	Item4	Item8	Item5	Item6	Item7	Item1	
	Item2	0.750 ^a	-0.195	-0.135	-0.035	-0.059	-0.048	-0.020	-0.031
	Item3	-0.195	0.779 ^a	-0.058	-0.069	-0.009	-0.180	-0.100	-0.016
	Item4	-0.135	-0.058	0.769 ^a	-0.081	-0.018	-0.078	0.011	-0.110
Anti-image	Item8	-0.035	-0.069	-0.081	0.670 ^a	0.017	-0.015	-0.412	0.004
Correlation	Item5	-0.059	-0.009	-0.018	0.017	0.758 ^a	-0.112	-0.051	-0.145
	Item6	-0.048	-0.180	-0.078	-0.015	-0.112	0.746 ^a	-0.296	-0.129
	Item7	-0.020	-0.100	0.011	-0.412	-0.051	-0.296	0.668 ^a	-0.054
	Item1	-0.031	-0.016	-0.110	0.004	-0.145	-0.129	-0.054	0.761 ^a

The extraction factor was carried out using the PCA method followed by factor rotation using the varimax method. Table 10 shows the value of communalities at item3, item4, and item6 less than 0.50. Variables should generally have communalities of greater than 0.50 to be retained in the analysis. Re-specification of factor analysis can include such option as (1) deleting a variable, (2) changing rotation methods, or (3) increasing or decreasing the number of factors (Hair, et. al. 2010). Therefore, an EFA test is carried out again by issuing items with the value of communalities below 0.50. EFA is carried out several times until all the prerequisites are significance.

Tabel 10
Communalities

Item	Item2	Item3	Item4	Item8	Item5	Item6	Item7	Item1
Initial	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Extraction	.609	.456	.427	.631	.581	.496	.721	.541

The last EFA that fulfilled significant aspects were item2, item5, item7, and item8. All MSA values and the value of communalities above 0.50. Total Variance Explained, component matrix and Component Transformation Matrix show that two factors formed honesty. The questionnaire reliability is 0.448. The generally agreed upon lower limit for Cronbach’s alpha is 0.70 (Robinson, et. al., 1991). Therefore, efforts need to be made to produce a reliable questionnaire.

The two factors can affect the reliability, namely (1) the length of the test and the quality of the instrument, (2) the condition during data collection (Retnawati, 2017). This study seeks to fix the quality of questionnaire items to improve reliability. After that, the instrument was tested. A total of 1261 questionnaires that have been filled by parents of students who carry out physics learning were collected.

The EFA results show the Chi-square value in the Bartlett test of 4059.11 with degrees of freedom 28 and p-less than 0.01 values. The KMO is 0.902 greater than 0.5. All MSA values above 0.50. All communalities values are above 0.50 except item4 (0.459) item1

(0.371), and item5 (0.154). EFA is done again with previously issued three items that have communalities below 0.50 so that the remaining five items.

The Chi-square value in the Bartlett test of 3048,892 with degrees of freedom 10 and p-values is less than 0.01. The KMO and Bartlett's Test value is 0.856. All MSA values and communalities are above 0.50. The total Cronbach's Alpha was 0.769. Nevertheless, the value of Cronbach's Alpha if Item Deleted on item6 is more than the total of Cronbach's Alpha (0.813). Therefore, item6 was dropped.

The EFA process is repeated again with previously issued item4, item1, item5, and item6. The Chi-square value in the Bartlett test is 1944.106 with degrees of freedom 6 and p-less than 0.01 values. The KMO is 0.816. All MSA and communalities values above 0.50 (Table 11 and Table 12).

Table 11
Anti-image Matrices

		Item2	Item3	Item7	Item8
Anti-image Correlation	Item2	.826 ^a	-.236	-.329	-.187
	Item3	-.236	.808 ^a	-.333	-.273
	Item7	-.329	-.333	.794 ^a	-.234
	Item8	-.187	-.273	-.234	.843 ^a

Table12
Communalities

Item	Item2	Item3	Item7	Item8
Initial	1.000	1.000	1.000	1.000
Extraction	.661	.699	.718	.630

Table 13 shows that one factor formed honesty. Only the factors having eigenvalues greater than 1 are considered significant (Hair, 2010). It reinforced by Table 14 and Figure 2 that there is only one factor forming the honesty variable which covers item2, item3, item7, and item8. The value of the loading factor above 0.50 is generally considered necessary for practical significance (Hair, 2010). The

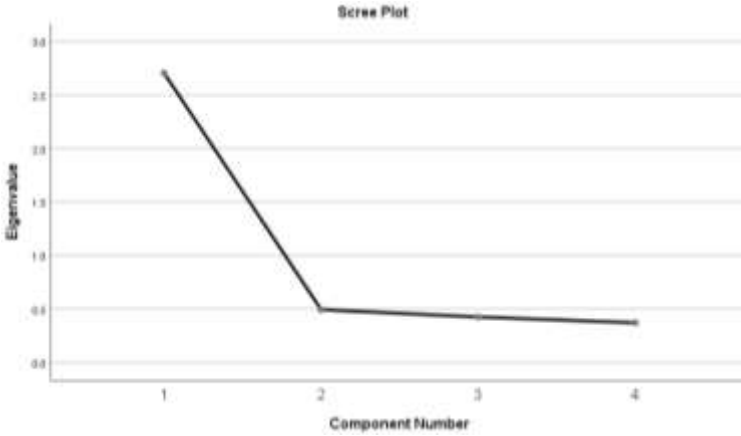
results of factor analysis crosscheck with the framework. It has consulted with experts to produce a variable forming factor honesty that fixes, namely maintaining the trust of other by not cheating.

Table 13
Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	2.707	67.673	67.673	2.707	67.673	67.673
2	.495	12.373	80.046			
3	.426	10.648	90.695			
4	.372	9.305	100.000			

Table 14
Component Matrix

	Component
	1
Item2	.813
Item3	.836
Item7	.847
Item8	.794



Picture 2
Scree plot factor of honesty

The questionnaire reliability is 0.840 (Table 15). The reliability distribution of each item is in Table 16. All items have Corrected Item-Total Correlation above 0.30. Rules of thumb suggest that the item-total correlation exceed 0.30 (Robinson, et. al., 1991). All the Cronbach's Alpha if item deleted is less than the total of Cronbach's Alpha (Table 16). Thus, the honesty questionnaire has fulfilled the reliability aspects. It was sufficient to measure honesty in the field.

Table 15
Questionnaire reliability

Cronbach's Alpha	N of Items
0.840	4

Table 16
Reliability of questionnaire items

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
Item2	10.2989	4.822	0.659	0.804
Item3	10.3545	4.628	0.692	0.789
Item7	10.1932	4.643	0.708	0.782
Item8	10.3092	4.756	0.635	0.815

Concluding (step 8)

Based on the results of expert judgment and EFA, a valid and reliable instrument is established to measure honesty through the learning process. The instruments include a questionnaire, observation sheet, and interview guideline sheet. The questionnaire consists of 4 items (Table 17). It formed from maintaining the trust of others by not cheating. Alpha Cronbach was 0.840. The observation sheet and interview guidelines consist of 3 items. The observation sheet was formed from not cheating and writing down real trial data. Interview

guidelines were formed from not lying, not cheating, and maintaining the trust of others.

Table 17
The final questionnaires

No	Statement
1	My child told someone else to do his job.
2	My child borrows other people's items but does not return.
3	My child asks for permission when he/she will use other people's items.
4	My child did not ask permission first when he/she wanted to use other people's items.

DISCUSSION

The purpose of this study is to produce an instrument to assess honesty that is valid and reliable. Some honesty theories are used to compile indicators. These indicators include not lying, not cheating, maintaining the trust of others, and writing down real trial data. The indicator is used as the basis for determining instruments, including questionnaires, observation sheets, and interview guidelines.

The indicators reflected in the questionnaire with eight items, interview guideline and observation with three items. All instruments are validated in terms of content by experts from the university academics and physics teacher. Instruments fulfill content validation if the results of expert judgment are good criteria. The interview sheet is an instrument that has undergone the most changes based on expert advice, especially from university academics. These changes have found in the question content. The expert recommends that questions are sharper in exploring student honesty.

The questionnaire also tested to determine to construct validity and reliability. The limited experiment was the most difficult stages. It has required more time than the specified time allocation. The majority of the students forgot to return the questionnaire. Some students rent boarding rooms, so they rarely met the parents. Some students also lived separately with their parents for many years

because of work outside the island and even abroad. The questionnaire was filled by the grandparents, uncle, or aunt.

EFA was conducted to determine the construct validity and Cronbach Alpha showing questionnaire reliability. In this study, EFA was carried out several times to produce significant data. The first EFA analyzed eight items from 1167 respondents' data. As a result, the value of KMO is 0.721. MSA value is above 0.50 for all items. However, not all commonalities values fulfill significance because they are below 0.50. These items are item3, item4, and item6. Therefore, the next EFA was carried out by previously issuing the three items. EFA stopped when significant results have been obtained. Significant EFA results are generated from item2, item8, item5, and item7. KMO is 0.541. All MSA values and commonalities are above 0.50. The four items fulfill construct validity. However, reliability is only 0.448. So the instrument is not reliable. Questionnaires with 8 items were repaired to have better quality. Here's an example.

Item3

My child does not return other people's items that he borrowed. (Before)

My child borrows other people's items but does not return. (After)

Item6

My child returns other borrowed goods in good condition and is clean. (Before)

My child returned the borrowed goods in good condition and clean. (After)

A total of 1261 respondents filled out the questionnaire. As a result, KMO is 0.902 and all items have MSA values above 0.50. However, item4, item1, and item5 have communality below 0.50. Thus, EFA is carried out on the remaining 5 items. The result is KMO of 0.856. All MSA and communality values are above 0.50. Thus, five items fulfill the construct validation aspect and can be used to find out the honesty forming factors. The reliability of the five items is 0.769. It is reliable, but the value of Cronbach's Alpha Deleted Item on item6

is more than the total of Cronbach's Alpha (0.813). Therefore, item 6 was dropped. EFA is done again on four items that fulfill, namely item2, item3, item7, and item8.

As a result, KMO is 0.816. All items have MSA and communality values above 0.50. Thus, four items fulfilled the construct validation aspect and can be used to find out the honesty forming factors. The questionnaire reliability is 0.840. All value of Cronbach's Alpha if Item Deleted not more than a total of Cronbach's Alpha. Thus, the four items meet the standards as questionnaire items to measure honesty.

The honesty indicators of the questionnaire in the framework and EFA results turned out differently. Based on EFA, not laying is not a factor forming honesty. Theoretically, 'not laying' is the main indicator of honesty (Tillman, 2004; Komalasari, 2012). The EFA results show that honesty is formed by the factor of maintaining the trust of others through not cheating.

Judging from the characteristics of respondents, they are generally not parents of students. Students generally come from families with low economic levels. Therefore, parents are forced to work outside the city and even abroad. Not even a few parents of students are divorced. Students live with grandparents, even grandmothers, or other families. Looking at the conditions, it is possible for respondents to be less careful in assessing student honesty. So, it is necessary for verifying by interviewing.

DePaulo, et.al (1996) examined lying behavior in a community. Some people consider lying as a process of everyday social interaction. They do not consider lies to be serious problems. So, they are not worried about getting caught up in the law. Lies in daily intimacy only result in relationships becoming less intimate than relationships without lies. Thus, it only affects the ethics of getting along. If so, the discussion of lying will lead to culture. That certainly requires further thought and in-depth research.

The development of an honest assessment instrument in the field of science has been carried out by Novodvorsky (1993). Apart from some of the differences, this research was necessary to be done

because times have changed. The living value determinants have increased. “Researchers need to provide reliable and valid evaluation tools that are not sustainable as they did in the past by developing instruments that will hardly meet the simple criteria (LaForgia, 1988: 407).

This instrument was expected to be used as a reference to measure the honesty, especially students in the process of learning at school. The condition of students in the family also needs attention. It aims to ensure that parents' assessment of students is not biased. Hopefully, the practice of dishonesty detected as early as possible.

Further research, researchers want to develop honesty instruments through a learning process in a comprehensive manner that covers the school, family and community environment. Communities in this context are people around the place where students live, such as neighbors. The participation of the community to assess the students' honesty is very important. The full-day school phenomenon causes students to learn honesty from family and school. As a result, students lack the provision to survive in a pluralistic society. Even though, education is expected to prepare students to live in society. Some experts also recommended it. Honesty measurement instruments will be able to assess honesty comprehensively.

LIMITATION

The contents of the questionnaire were not checked by interviewing the parents of students. It is necessary for verifying. For researchers who want to use this instrument, it is recommended to conduct interviews with students' parents as a means of crossing the questionnaire data.

CONCLUSION

The valid and reliable honesty assessment instruments were produced, namely questionnaires, observation sheets, and interview sheet. All instruments fulfilled the content validity based on expert judgment assessment. EFA shows that the questionnaire fulfilled the construct validity. It formed by maintaining the trust of others by not cheating and consisted of four items. The reliability has fulfilled with a coefficient of 0.840.

SUGGESTION

The results of this study need to be followed up with data collection in the field and dissemination. Then, ten steps to developing an honesty instrument are fulfilled.

BAB IV

VALIDITAS INSTRUMEN

A. Pengertian

Sebagian kalangan sering menyebut validitas dengan kesahihan. Validitas berasal dari kata *validity* yaitu sejauh mana ketepatan dan kecermatan alat ukur dalam melakukan fungsi ukurnya. Kerlinger (1986) mendefinisikan validitas sebagai sebuah kriteria untuk mengetahui sejauhmana mana suatu alat ukur mampu mengukur apa yang seharusnya diukur. Linn & Gronlund (1995) menjelaskan validitas mengacu pada kecukupan dan kelayakan interpretasi dari sebuah penilaian. Messick (1989) memandang validitas sebagai sebuah kebijakan evaluatif yang terintegrasi tentang sejauhmana fakta empiris dan alasan teoretis mendukung kecukupan dan kesesuaian inferensi dan tindakan berdasarkan skor instrumen. Sudaryono, Margono, dan Rahayu (2013) menyatakan validitas sebagai suatu konsep yang berkaitan dengan sejauhmana alat ukur telah mengukur apa yang seharusnya diukur. Adapun Retnawati (2017) menjelaskan validitas sebagai ukuran yang menunjukkan dukungan fakta empiris dan alasan teoretis terhadap interpretasi skor suatu instrument dan terkait dengan kecermatan pengukuran. Berdasarkan beberapa definisi validasi oleh para ahli, dapat disimpulkan bahwa validitas adalah suatu standar kualitas instrumen dalam mengukur sesuatu yang seharusnya diukur secara cermat.

Sudaryono, Margono, dan Rahayu (2013) membagi validitas instrumen menjadi 2, yaitu validitas logis dan validitas empiris. Validitas logis merupakan ukuran validitas yang mengukur berfungsi tidaknya suatu instrument berdasarkan kriteria yang telah ditentukan meliputi kriteria materi, konstruksi, dan bahasa. Suatu instrumen dikatakan memiliki validitas tinggi jika instrumen tersebut menjalankan fungsi ukurnya. Artinya, hasil pengukuran mencerminkan secara tepat fakta atau keadaan sesungguhnya dari apa

yang diukur. Hal ini juga didukung oleh pendapat Djaali (2004) bahwa validitas satu instrumen harus selalu dikaitkan dengan tujuan atau pengambilan keputusan tertentu. Adapun validitas empiris adalah merupakan validitas instrumen yang didasarkan pada hasil ukur instrumen yang bersangkutan, baik melalui ujicoba maupun pengukuran yang sesungguhnya.

Ahli lain yaitu Nunnally (1978), Allen & Yen (1979), Fernandes (1984), Woolfolk & McCane (1984), Kerlinger (1986), dan Lawrence (1994) dalam Heri Retnawati (2013) mengelompokkan validitas menjadi tiga tipe, yaitu: (1) validitas isi, (2) validitas konstruk, dan (3) validitas kriteria (*criterion-related*). Validitas isi merupakan analisis rasional terhadap domain yang hendak diukur untuk mengetahui keterwakilan instrumen dengan kemampuan yang hendak diukur. Validitas konstruk adalah validitas yang menunjukkan sejauhmana instrumen mengungkap suatu kemampuan atau konstruk teoretis tertentu yang hendak diukur. Validitas berdasarkan kriteria dimaksudkan untuk menjawab pertanyaan sejauh mana tes memprediksi kemampuan peserta di masa mendatang (*predictive validity*) atau mengestimasi kemampuan dengan alat ukur lain dengan tenggang waktu yang hampir bersamaan (*concurrent validity*).

Pada bab ini hanya akan dibahas secara lebih mendalam tentang validitas isi dan validitas konstruk beserta pembuktiannya. Pembahasan terkait validitas empiris terdapat pada Bab 4.

B. Validitas Isi dan Cara Pembuktiannya

Validitas isi menunjukkan sejauh mana pertanyaan, tugas atau butir instrumen mampu mewakili secara keseluruhan dan proporsional keseluruhan perilaku sampel yang menjadi tujuan pembelajaran yang akan diukur pencapaiannya (Gregory, 2000) dalam Retnawati (2013) mengungkapkan bahwa validitas isi terkait dengan analisis rasional terhadap domain yang hendak diukur untuk mengetahui keterwakilan instrumen dengan kemampuan yang hendak diukur.

Sudaryono, Margono, dan Rahayu (2013) menyatakan bahwa penentuan validitas isi suatu instrumen dapat dilakukan melalui penelaah kisi-kisi instrumen untuk memastikan bahwa item instrumen

sudah mewakili atau mencerminkan keseluruhan konten atau materi yang seharusnya dikuasai secara proporsional. Oleh karena itu, validitas isi tidak mempunyai besaran tertentu yang dihitung secara statistika, tetapi dipahami bahwa instrument itu sudah valid berdasarkan telaah kisi-kisi instrument. Mardapi (2016) menjelaskan lebih lanjut bahwa validitas isi data dibuktikan melalui 2 tahap. Pertama, yaitu kesesuaian antara indikator pada kisi-kisi dengan definisi operasional dari konstruk instrument. Kedua yaitu menentukan kesesuaian antara indikator dengan item instrumen.

Operasional penentuan validitas isi dilakukan melalui kesepakatan ahli bidang studi atau sering disebut dengan *domain* (Retnowati, 2013). Sebuah angket dibuktikan memenuhi validitas isi jika ahli (*expert*) meyakini bahwa instrumen tersebut mengukur penguasaan kemampuan yang didefinisikan dalam domain ataupun juga konstruk psikologi yang diukur. Caranya dengan menyelenggarakan diskusi panel (Sumaryono, Margono, Rahayu, 2013) atau FGD (*Forum Group Discussion*) Mardapi (2017) yaitu berkumpulnya ahli atau pakar yang dipandang memiliki keahlian yang ada hubungannya dengan suatu bahasan yang diujikan untuk memberikan pendapat dan rekomendasi terhadap isi atau materi dalam suatu instrument penelitian. Saat ini FGD telah banyak dipilih peneliti untuk melakukan validasi isi karena membutuhkan waktu yang relatif singkat.

Pengujian validitas isi memerlukan alat uji guna memastikan hasilnya. Retnowati (2013) dan Mardapi (2017) menyatakan 2 (dua) alat uji validitas isi yaitu (1) indeks validitas yang diusulkan oleh Aiken (1985) dan (2) indeks kesepakatan ahli yang disarankan oleh Gregory (2007).

Indeks Aiken V merupakan indeks kesepakatan rater terhadap kesesuaian butir (atau sesuai tidaknya butir) dengan indikator yang ingin diukur menggunakan butir tersebut melalui persamaan (1).

$$V = \frac{\sum s}{n(c - 1)}$$

Keterangan:

- V = indeks kesepakatan rater mengenai validitas butir
- s = skor yang ditetapkan setiap rater dikurangi skor terendah dalam kategori yang dipakai ($s = r - lo$, di mana r = skor kategori pilihan rater dan lo = skor terendah dalam kategori penyekoran);
- n = banyaknya rater
- c = banyaknya kategori yang dapat dipilih rater.

Indek V menunjukkan kategori validasi isi suatu instrumen. Jika indeks V kurang atau sama dengan 0,4 dikatakan validitas kurang, 0,4-0,8 dikatakan validitas sedang, dan jika lebih besar dari 0,8 dikatakan sangat valid. Contoh perhitungan validasi isi menggunakan V Aiken terdapat pada

Contoh 4.1

Formula penentuan indeks kesepakatan ahli oleh Gregory (2007) ditunjukkan dengan pemberian skor antara 0-1. Hal ini dilakukan dengan sebelumnya membuat tabel kontingensi pada dua ahli. *Kategori pertama*, yaitu tidak relevan dan kurang relevan diubah menjadi kategori relevansi lemah. *Kategori kedua*, yaitu cukup relevan dan sangat relevan diubah menjadi kategori baru yaitu relevansi kuat. Indeks kesepakatan ahli untuk validitas isi merupakan perbandingan banyaknya butir dari kedua ahli dengan kategori relevansi kuat dengan keseluruhan butir. Penentuan kriteria hasil validitas isi menggunakan kesepakatan ahli oleh Gregory (2007) sama dengan indeks V Aiken. Contoh perhitungan validasi isi menggunakan V Aiken terdapat pada Contoh 4.2.

Lebih jauh Retnowati (2013) menjelaskan langkah-langkah membuktikan validitas isi.

1. Memberikan kisi-kisi dan butir instrumen, disertai rubrik penskorannya jika ada kepada beberapa ahli yang sesuai dengan bidang yang diteliti untuk meminta masukan. Banyaknya ahli paling tidak 3 orang ahli dengan kepakaran yang relevan dengan bidang yang diteliti.
2. Masukan yang diharapkan dari ahli berupa kesesuaian komponen instrumen dengan indikator, indikator dengan butir, benarnya substansi butir, kejelasan kalimat dalam butir, kalimat-kalimat tidak membingungkan, format tulisan, simbol, dan gambar yang cukup jelas. Proses ini sering disebut telaah kualitatif yang meliputi aspek substansi, bahasa, dan budaya.
3. Berdasarkan masukan ahli tersebut, kisi-kisi dan atau instrumen kemudian diperbaiki.
4. Meminta ahli untuk menilai validitas butir, berupa kesesuaian antara butir dengan indikator. Penilaian ini dapat dilakukan misalnya dengan skala Likert (skor 1: tidak valid, skor 2 = kurang valid, skor 3 = cukup valid, skor 4 = valid, skor 5 = sangat valid). Dapat pula penskoran dengan melihat relevansi butir dengan indikator (skor 1: tidak relevan, skor 2 = kurang relevan, skor 3 = cukup relevan, skor 4= relevan, skor 5 = sangat relevan).
5. Menghitung indeks kesepakatan ahli (*rater agreement*) dengan indeks Aiken V atau indeks Gregory, yang merupakan indeks untuk menunjukkan kesepakatan hasil penilaian para ahli tentang validitas, baik untuk butir maupun untuk perangkatnya.

Contoh 4.2 Penentuan validasi menggunakan formulas V Aiken

Data diambil dari aspek kejujuran dari BAB III.

Tabel 4.1 Angket kejujuran

No	Pernyataan	Skor			
		1	2	3	4
1	Anak saya bicara berdasarkan fakta di lapangan				
2	Anak saya menyuruh orang lain untuk mengerjakan tugasnya				
3	Anak saya tidak mengembalikan barang orang lain yang dipinjamnya				
4	Apa yang dibicarakan anak saya berbeda dengan fakta				
5	Anak saya mengerjakan tugasnya sendiri				
6	Anak saya mengembalikan barang orang lain yang dipinjam dalam keadaan baik dan bersih				
7	Anak saya minta izin ketika mau memakai barang orang lain				
8	Anak saya tidak meminta izin terlebih dahulu saat mau memakai barang orang lain				
	Skor total				

Lembar penilaian ahli tersaji berikut ini.

Tabel 4.2 Lembar penilaian validitas isi oleh ahli

No. butir	Hasil penilaian ahli dengan memberi centang				
	1	2	3	4	5
	Tidak relevan	Kurang relevan	Cukup relevan	Relevan	Sangat relevan
1					
2					
Dst					

Berikut hasil penilaian oleh ahli

Tabel 4.3 Skor penilaian oleh 3 validator

No butir	Validator 1	Validator 2	Validator 3
1	5	4	4
2	5	5	5
3	5	5	5
4	5	4	5
5	5	5	5
6	5	5	5
7	5	5	5
8	5	5	5

Proses penghitungan V Aiken untuk masing-masing butir angket.

Tabel 4.4 V Aiken masing-masing butir angket

No butir	Rater1	Rater 2	Rater 3	s1	s2	s3	Σs	V
1	5	4	4	4	3	3	10	0,83
2	5	5	5	4	4	4	12	1,00
3	5	5	5	4	4	4	12	1,00
4	5	4	5	4	3	4	11	1,00
5	5	5	5	4	4	4	12	0,92
6	5	5	5	4	4	4	12	1,00
7	5	5	5	4	4	4	12	1,00
8	5	5	5	4	4	4	12	1,00

Untuk keseluruhan skala, diketahui besar indeks V Aiken adalah sebagai berikut.

Tabel 4.5 Indeks V Aiken perangkat

Skala	Rater1	Rater 2	Rater 3	S1	S2	S3	Σs	V
Butir 1-8	40	38	39	32	30	31	93	0,79

Berdasarkan Tabel...., maka dikatakan bahwa perangkat memiliki indeks V Aiken 0,79 dengan kategori sedang.

Penerapan Isu Lingkungan 4.1

Penggunaan V Aiken sebagai metode penguji validasi isi dapat dilihat pada artikel berjudul: “*The Effect of Laboratory Work Style and Reasoning with Arduino to Improve Scientific Attitude*” yang dapat dilihat melalui web: <https://eric.ed.gov/?id=EJ1210983>

Contoh 4.2

Penentuan validasi menggunakan formulas Gregory (2007)

Langkah awal yaitu dengan menyiapkan lembar penilaian validator

Tabel 4.6 Lembar penilaian validator

No butir	Hasil penilaian ahli dengan memebrikan tanda centang (v)			
	1	2	3	4
	Tidak relevan	Kurang relevan	Cukup relevan	Relevan
1				
2				
3				
4				
5				
6				
7				
8				

Ahli kemudian menilai kesesuaian butir dengan indikator. Hasilnya kemudian dimasukkan pada tabel berikut.

Tabel 4.7 Hasil tabulasi 2 rater

No butir	Ahli 1	Ahli 2
1	4	3
2	3	4
3	4	4
4	2	3
5	4	4
6	2	1
7	4	4
8	3	4

Hasil penilaian ahli kemudian diubah kategorinya, yaitu kategori tidak relevan (1) dan kurang relevan (2) menjadi kategori lemah. Adapun kategori cukup relevan (3) dan relevan (4) diubah ke kategori kuat. Tabel ... kemudian berubah sebagai berikut.

Tabel 4.8 Hasil kategori ulang dua ahli

No butir	Ahli 1	Ahli 2
1	Kuat	Kuat
2	Kuat	Kuat
3	Kuat	Kuat
4	Lemah	Kuat
5	Kuat	Kuat
6	Lemah	Lemah
7	Kuat	Kuat
8	Kuat	Kuat

Berikut tabel kontingensi ulang dari dua ahli

Tabel 4.9 Kontingensi ulang dari dua ahli

		Rater 1	
		Lemah	Kuat
Rater 2	Lemah	1	0
	Kuat	1	6

Retnowati (2013) menyebutkan bahwa, indeks kesepakatan ahli untuk validitas isi merupakan perbandingan banyaknya butir dari kedua ahli dengan kategori relevansi kuat dengan keseluruhan butir, sehingga dihasilkan sebagai berikut.

$$Koefisien\ validitas\ isi = \frac{6}{(1 + 1 + 0 + 6)} = 0,75$$

Mengikuti ketentuan kriteria validitas isi pada indeks V Aiken, maka validitas isi instrumen dengan menggunakan indeks kesepakatan ahli melalui formula Gregory (2007) menunjukkan nilai 0,75 dengan kategori sedang.

Selain penggunaan indeks V Aiken dan kriteria Gregory, penentuan validitas isi juga dapat dilakukan menggunakan kriteria (Mardapi, 2008) seperti contoh 3.

Contoh 4.3 Penentuan validasi isi menggunakan kriteria

Instrumen yang digunakan adalah angket kejujuran (Sarah, Prasetyo, Wilujeng, 2018) yang sebelumnya sudah dijelaskan pada bab 1 dan 2 untuk dinilai aspek validitas isi. Pengujian validasi isi dilakukan oleh 3 pakar yang membidangi proses pembelajaran dari akademisi perguruan tinggi. Semua ahli menilai instrumen berdasarkan 3 aspek pengukuran instrumen, yaitu isi, bahasa, dan konstruksi yang masing-masing aspek teruraikan dalam item penilaian angket seperti pada Tabel 4.10

Tabel 4.10 Dokumen penilaian item angket

No	Aspek	Item penilaian angket	Nomor item
1	isi	Kesesuaian dengan variabel yang akan dicapai	1
2	Bahasa	Penggunaan struktur kalimat yang sederhana	2
		Penggunaan bahasa baku	3
		Penggunaan struktur kalimat yang efektif	4
		Penggunaan bahasa mudah dipahami	5
		Penggunaan istilah tidak menimbulkan makna ganda	6
3	Konstruksi	Kejelasan identitas angket	7
		Kejelasan petunjuk pengisian angket	8
		Kejelasan pernyataan yang dirumuskan	9

Para ahli diminta memberikan masukan atas angket kejujuran dengan cara memberikan tanda centang (v) pada kolom yang menunjukkan hasil penilaian ahli pada skala 1 hingga 5. Adapun keterangan masing-masing skala, yaitu 1 = sangat kurang, 2 = kurang, 3 = cukup, 4 = baik, dan 5 = sangat baik seperti terlihat pada Tabel 4.11 berikut.

Tabel 4.11 Lembar validasi angket kejujuran oleh ahli

No	Item penilaian angket	Skor					Saran
		1	2	3	4	5	
1	Kesesuaian dengan variabel yang akan dicapai						
2	Penggunaan struktur kalimat yang sederhana						
	Penggunaan bahasa baku						
	Penggunaan struktur kalimat yang efektif						
	Penggunaan bahasa mudah dipahami						
	Penggunaan istilah tidak menimbulkan makna ganda						
3	Kejelasan identitas angket						
	Kejelasan petunjuk pengisian angket						
	Kejelasan pernyataan yang dirumuskan						
Jumlah							

Hasil penilaian masing-masing item oleh ahli dicari rata-ratanya pada setiap aspek, yaitu isi, bahasa, dan konstruksi. Berikut hasilnya.

Tabel 4.13 Hasil uji validasi angket kejujura3 (tiga) ahli

Aspek	Item penilaian	Nomor	Ahli			Rata-rata skor ahli dan kriteria
			1	2	3	
isi	Kesesuaian dengan <i>living values</i> yang akan dicapai	1	5	4	4	4,33
Bahasa	Penggunaan struktur kalimat yang sederhana	2	5	5	5	24,67

	Penggunaan bahasa baku	3	5	5	5	
	Penggunaan struktur kalimat yang efektif	4	5	4	5	
	Penggunaan bahasa mudah dipahami	5	5	5	5	
	Penggunaan istilah tidak menimbulkan makna ganda	6	5	5	5	
Konstruksi	Kejelasan identitas	7	5	5	5	14,67
	Kejelasan petunjuk pengisian instrumen	8	5	5	5	
	Kejelasan pernyataan yang dirumuskan	9	5	4	5	

Tahapan selanjutnya, adalah memberikan kriteria apakah angket memenuhi validitas isi atau tidak berdasarkan tabel konversi (Mardapi, 2008).

Tabel 4.14 Konversi hasil pengukuran

No	Interval skor	Category
1	$x \geq \bar{x} + 1.SBx$	Sangat baik
2	$\bar{x} + 1.SBx > x \geq \bar{x}$	Baik
4	$\bar{x} > x \geq \bar{x} - 1.SBx$	Tidak baik
5	$x < \bar{x} - 1.SBx$	Sangat tidak baik

Keterangan:

\bar{x} = mean ideal = $\frac{1}{2}$ (skor tertinggi ideal + skor terendah ideal)

SBx = simpangan baku skor ideal (SB_i) = $\frac{1}{6}$ (skor tertinggi ideal - skor terendah ideal)

x = skor yang dicapai

(Mardapi, 2008)

Instrumen dikatakan memenuhi validitas isi oleh ahli jika memenuhi kriteria “sangat baik” dan “baik”.

Berdasarkan tabel konversi, maka dilakukan penyesuaian konversi pada masing-masing aspek, seperti terlihat pada Tabel 4.15, Tabel 4.16, dan Tabel 4.17.

Tabel 4.15 Konversi hasil pengukuran angket aspek isi

No	Rentang skor	Nilai	Kategori
1	$x > 4,20$	A	sangat baik
2	$3,40 < X \leq 4,20$	B	baik
3	$2,60 < X \leq 3,40$	C	cukup baik
4	$1,80 < X \leq 2,60$	D	kurang baik
5	$X \leq 1,80$	E	sangat kurang baik

Tabel Tabel 4.16 konversi hasil pengukuran angket aspek bahasa

No	Rentang skor	Nilai	Kategori
1	$x > 21,00$	A	sangat baik
2	$17,00 < X \leq 21,00$	B	baik
3	$13,00 < X \leq 17,00$	C	cukup baik
4	$9,00 < X \leq 13,00$	D	kurang baik
5	$X \leq 9,00$	E	sangat kurang baik

Tabel Tabel 4.17 konversi hasil pengukuran angket aspek konstruksi

No	Rentang skor	Nilai	Kategori
1	$x > 12,60$	A	sangat baik
2	$10,20 < X \leq 12,60$	B	baik
3	$7,80 < X \leq 10,20$	C	cukup baik
4	$5,40 < X \leq 7,80$	D	kurang baik
5	$X \leq 5,40$	E	sangat kurang baik

Menggunakan hasil konversi, maka dapat diketahui kriteria validasi isi dari masing-masing aspek, yaitu seperti terlihat pada Tabel 4.18.

Tabel 4.18 Hasil validasi angket oleh ahli

Aspek	Item penilaian	Ahli			Skor maksimum	Rata-rata skor ahli dan kriteria
		1	2	3		
isi	Kesesuaian dengan <i>living values</i> yang akan dicapai	5	4	4	5	4,33 (sangat baik)
Bahasa	Penggunaan struktur kalimat yang sederhana	5	5	5	25	24,67 (sangat baik)
	Penggunaan bahasa baku	5	5	5		
	Penggunaan struktur kalimat yang efektif	5	4	5		
	Penggunaan bahasa mudah dipahami	5	5	5		
	Penggunaan istilah tidak menimbulkan makna ganda	5	5	5		
Konstruksi	Kejelasan identitas	5	5	5	15	14,67 (sangat baik)
	Kejelasan petunjuk pengisian instrumen	5	5	5		
	Kejelasan pernyataan yang dirumuskan	5	4	5		

Karena skor angket pada semua aspek menunjukkan kriteria sangat baik, maka angket telah memenuhi validitas isi.

C. Validitas Konstruk dan Cara Pembuktiannya

Validitas konstruk (*construct validity*) adalah pengujian validitas yang dilakukan dengan melihat kesesuaian konstruksi butir yang ditulis dengan kisi-kisinya (Purwanto, 2010). Sudaryono, Margono, dan Rahayu (2013) menyatakan sebuah instrumen dikatakan memenuhi validitas konstruk jika telah cocok dengan konstruksi teoritik dimana instrumen itu dibuat. Validitas konstruk biasa digunakan untuk mengukur validitas-validitas konsep, baik yang sifatnya performasi tipikal seperti instrumen untuk mengukur sikap,

minat konsep diri, lokus kontrol, gaya kepemimpinan, motivasi berprestasi, dan lain, maupun yang sifatnya performansi maksimum seperti instrumen untuk mengukur bakat (tes bakat), intelegensi (kecerdasan intelektual), kecerdasan emosional dan lainnya.

Pengujian validitas konstruk menurut Retnowati (2017) dan Purwanto (2010) dapat dilakukan dengan cara menelaah kesesuaian butir instrumen dengan kisi-kisi dalam hal konstruksinya. Adapun variabel dan butir-butir instrumen telah direncanakan di dalam kisi-kisi. Variabel itu sendiri tersusun dari beberapa faktor. Sebuah faktor dapat diuraikan menjadi beberapa butir. Kegiatan menelaah butir instrumen dilakukan dengan mencermati kesesuaian penempatan butir-butir dalam faktornya. Jadi, hal ini berbeda dengan validitas isi. Jika validitas isi melihat kesesuaian butir dengan kisi-kisi dalam hal muatannya, penelaahan butir pada uji validitas konstruk dalam hal konstruksinya. Butir-butir dinyatakan valid jika menempati faktor yang sebelumnya sudah ada dan direncanakan dalam kisi-kisi.

Analisis faktor merupakan salah satu metode yang dapat digunakan dalam menguji validitas konstruk. Analisis faktor menjadikan kerumitan suatu hal menjadi lebih sederhana sehingga mudah dijelaskan. Analisis faktor itu sendiri merupakan salah satu analisis multivariat yang dirancang untuk meneliti sifat hubungan antar variabel dalam satu perangkat tertentu yang pada dasarnya menentukan pola hubungan tertentu (Purwanto, 2010). Retnowati (2017) menambahkan bahwa analisis faktor membuktikan validitas konstruk dengan cara membuktikan kebermaknaan skor hasil pengukuran (*meaningfulness*). Hal ini dilakukan dengan cara membuktikan bahwa konstruk instrumen memang ada (*exists*) dan kemudian dibuktikan hasil pengukurannya secara empiris (Retnowati, 2017). Metode lain untuk menguji validitas konstruk adalah (1) menelaah butir, (2) meminta pertimbangan ahli, (3) konvergensi dan diskriminabilitas, dan (4) multitrait-multimethod (MTMM). Pada bahasan ini hanya akan mengkaji validitas konstruk menggunakan analisis faktor.

Analisis faktor terbagi menjadi 2, yaitu analisis faktor eksploratori (*exploratory factor analysis, EFA*) dan analisis faktor konfirmatori (*confirmatory factor analysis, CFA*). Analisis faktor eksploratori bersifat mengeksplor data empiris untuk menemukan dan mendeteksi karakteristik dan hubungan antar variabel tanpa menentukan model pada data. Hasil analisis yang menjelaskan hubungan antar variabel semata tidak didasarkan pada teori yang ada (Retnawati, 2017). Purwanto (2010) menjelaskan lebih jauh lagi bahwa EFA digunakan untuk mengetahui faktor-faktor yang melandasi sehimpunan variabel atau ukuran. EFA tidak menghipotesiskan adanya sejumlah faktor dari butir-butir pengukur variabel. Artinya, butir-butir dibiarkan membentuk polanya sendiri dan menginformasikan ditemukan faktor-faktor.

Berbeda dengan EFA, CFA digunakan untuk menguji model yang telah diasumsikan. Kerlinger (1996) memperjelas bahwa CFA digunakan untuk menguji hipotesis mengenai struktur faktor dalam sehimpunan data. Garson (2006) juga menyampaikan bahwa CFA dibangun dengan mendasarkan pada informasi tentang struktur data dalam bentuk teori khusus atau hipotesis. Adapun teori khusus atau hipotesis yang dibangun didasarkan pada teori yang telah ada atau hasil penelitian sebelumnya. Jika pada EFA konstruk instrumen masih dicari ataupun dilakukan eksplorasi, sebaliknya pada CFA, model pengukuran telah ada teorinya. Konstruk instrumen hanya tinggal dibuktikan atau dikonfirmasi. Pembuktian tersebut dapat dilakukan salah satunya menggunakan *first order CFA*. Pada bahasan ini, hanya akan ditunjukkan contoh proses analisis menggunakan EFA. Berikut beberapa istilah teknis yang digunakan dalam analisis faktor, yaitu: (1) variabel atau butir, (2) faktor, (3) ekstraksi, (4) *eigenvalues*, (5) rotasi, (6) total common variance atau *communalities*, dan (7) *factor loadings*.

Variabel atau butir adalah unit terkecil pengukur yang diperkirakan akan mengukur satu dimensi atau variabel. Analisis faktor akan menguji apakah setiap butir itu memang mengukur masing-masing satu dimensi atau beberapa butir sebenarnya mengukur dimensi yang sama. Faktor adalah beberapa butir atau variabel yang

sebenarnya mengukur dimensi yang sama. Banyaknya butir/variabel yang digunakan dalam instrumen menjadikan peneliti sulit memahami dimensi yang dimiliki variabel. Analisis faktor akan membantu dalam menyederhanakannya. Arti dari ekstraksi, yaitu bahwa analisis faktor melakukan prosedur ekstraksi untuk memeras banyak butir hingga ditemukan beberapa faktor saja yang menjadi sarinya. *Eigenvalue* menyatakan ukuran nilai tertentu dari varians butir agar dapat dikonstruksi menjadi beberapa faktor. Setiap butir memiliki *eigenvalues* 1,00 bila digunakan untuk mengukur variabel. Sebuah faktor dimaksudkan untuk mengukur butir-butir yang mempunyai dimensi sama. Pekerjaan meringkas bekerja dengan baik jika faktor berperan lebih baik dibandingkan butir-butir berdiri sendiri, sehingga faktor harus mempunyai faktor di atas 1,000. Jadi, faktor hanya dipertahankan bila memiliki *eigenvalue* di atas 1,00. Rotasi pada analisis faktor dikamsudkans ebagai proses memutar sumbu mendekati koordinat butir sehingga diketahui pengelompokkan dan sumbangan butir kepada faktor. Varians umum (*common variance*) merupakan bagian dari varians butir tersebut yang disumbangkan kepada setiap faktor. Total varians umum yang disumbangkan pada semua faktor disebut komunalitas. Bila metode ekstraksi menggunakan *principal component analysis*, maka komunalitas akan selalu 1,00 (satu) karena metode ekstraksi ini tidak memisahkan varians umum, spesifik, dan galat. Adapun *factor loading* merupakan sumbangan varians bersama sebuah butir kepada faktor. Sebuah butir menjadi muatan sebuah faktor jika memberikan sumbangan yang besar hanya pada satu faktor tersebut dan tidak pada faktor lainnya (Purwanto, 2010).

Prosedur pelaksanaan analisis faktor meliputi: (1) menguji kelayakan analisis, (2) menyajikan matriks korelasi, (3) melakukan ekstraksi, (4) melakukan rotasi, dan (5) memberikan penamaan faktor. Berikut akan dijelaskan satu per satu (Purwanto, 2010).

Pengujian kelayakan analisis dilakukan untuk melihat terpenuhinya asumsi sebagai syarat dilakukannya analisis faktor. Kriteria tersebut memuat 2 hal. Pertama, yaitu harga koefisien *Keiser-Meyer-Olkin* (KMO) *sampling adequacy* untuk melihat kecukupan

sampel yang dianalisis. Kedua, yaitu Bartlett's *sphericity tes* atau uji Bartlett untuk melihat normalitas data yang akan dianalisis. Jika kedua asumsi terpenuhi, maka dapat disimpulkan bahwa analisis faktor memiliki kemampuan untuk generalisasi.

Penyajikan matriks korelasi memberikan informasi tentang korelasi antar butir. Butir yang saling berkorelasi tinggi menunjukkan pengukuran dimensi yang sama. Ekstraksi dilakukan untuk mendapatkan lebih sedikit faktor (*eigenvalue factor*) dari sejumlah banyak butir/variabel dan sumbangan faktor terhadap keseluruhan butir (*total variance explained*). Adapun rotasi bertujuan untuk melakukan proses yang belum dilakukan oleh prosedur ekstraksi dengan menarik butir-butir ke dalam faktor-faktor terdekat. Beberapa metode rotasi seperti varimax, quartimax, equamax, dan oblimin. Adapun langkah terakhir yaitu menamai (*labeling*) faktor yang terbentuk dari proses ekstraksi dan rotasi dengan prinsip kesamaan ciri butir yang menjadi muatan faktor.

Contoh 4.4

Perhitungan validitas konstruk angket *living values* (kejujuran) dalam proses pembelajaran menggunakan EFA dengan SPSS versi 25.00

Uji validitas konstruk dilakukan dengan sebelumnya melakukan ujicoba angket yang sebelumnya telah memenuhi validasi isi oleh ahli (*expert judgement*). Angket disebar ke orang tua siswa SMA/MA di Kabupaten Wonosobo melalui bantuan guru fisika. SMA/MA yang terlibat dalam penelitian berjumlah 9 terdiri dari 7 SMA/MA negeri dan 2 SMA/MA swasta. Angket yang terkumpul sebanyak 1261 yang memenuhi syarat untuk dilakukan analisis. Data kemudian dianalisis menggunakan EFA yang bertujuan mengetahui faktor yang mengkonstruksi kejujuran sekaligus membuktikan validitas konstruk angket. Analisis ini menggunakan SPSS versi 25.00.

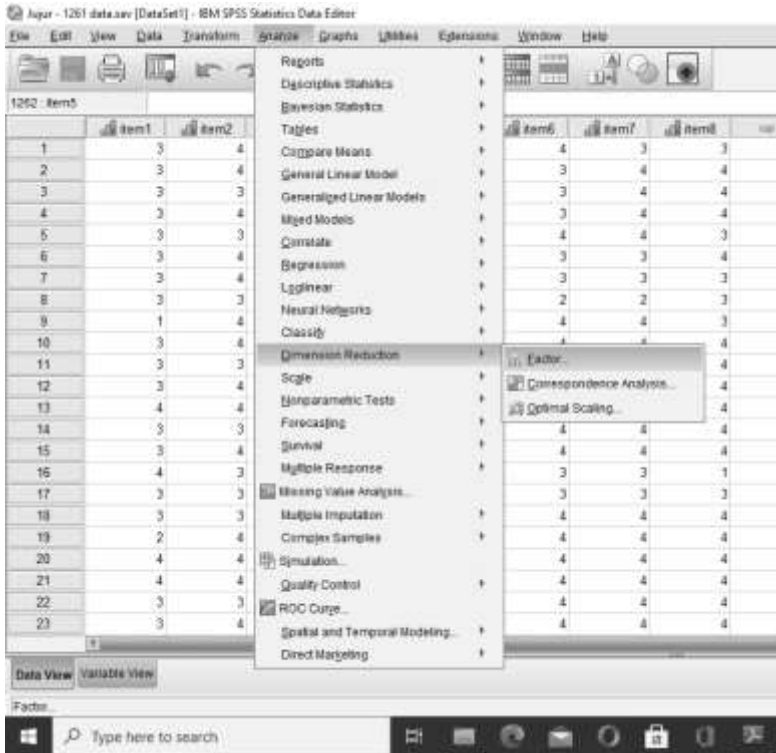
Berikut langkah analisis menggunakan SPSS dan cara penyelesaiannya.

Menyiapkan data dalam file SPSS pada data view, dengan variabel merupakan butir-butir instrumen pada variabel view.

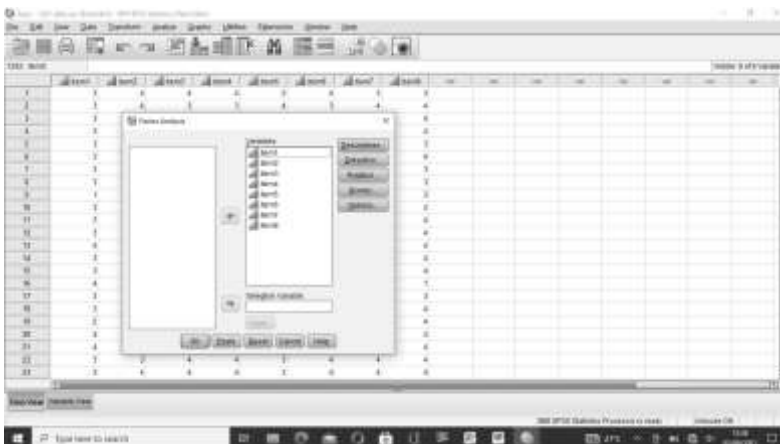
	Item1	Item2	Item3	Item4	Item5	Item6	Item7	Item8
1	3	4	4	4	2	4	3	3
2	3	4	3	3	4	3	4	4
3	3	3	4	3	3	3	4	4
4	3	4	4	4	3	3	4	4
5	3	3	4	4	3	4	4	3
6	3	4	4	3	3	3	3	4
7	3	4	3	4	3	3	3	3
8	3	3	4	4	4	2	2	3
9	1	4	3	2	4	4	4	3
10	3	4	4	4	4	4	4	4
11	3	3	4	4	3	4	4	4
12	3	4	4	4	4	4	4	4
13	4	4	4	4	3	3	3	4
14	3	3	4	2	4	4	4	4
16	3	4	4	4	4	4	4	4
16	4	3	3	2	3	3	3	1
17	3	3	3	2	2	3	3	3
18	3	3	4	3	3	4	4	4
19	2	4	4	3	3	4	4	4
20	4	4	4	4	4	4	4	4
21	4	4	4	3	3	4	4	4
22	3	3	4	4	3	4	4	4
23	3	4	4	4	3	4	4	4

Gambar Data kejujuran 1261 responden

Klik **analyze** → **Factor** sehingga muncul **Factor Analysis**

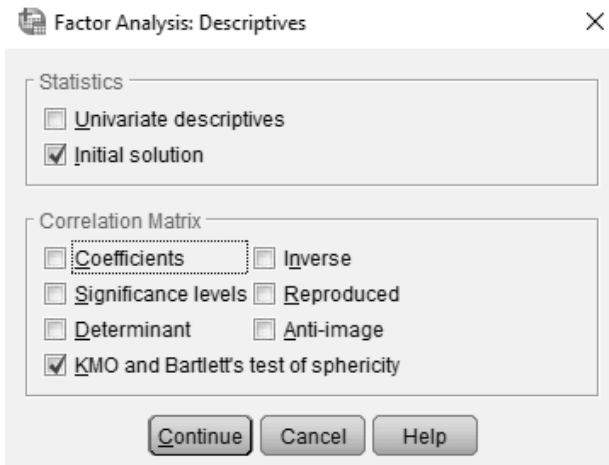


Gambar tahapan analisis faktor



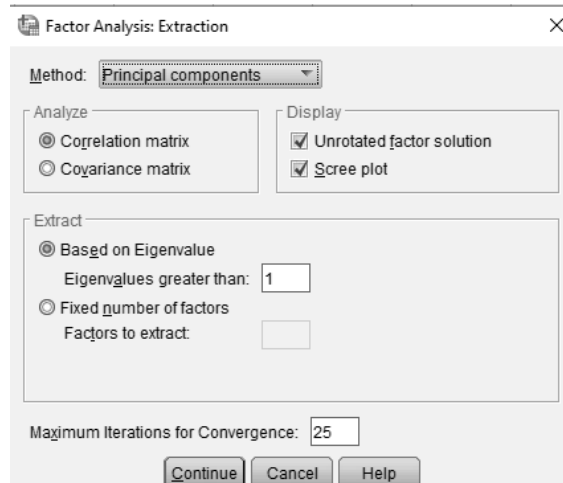
Gambar proses memasukkan data untuk analisis faktor

Klik **Descriptive** → Klik output yang dikehendaki, misalnya **KMO and Bartlett's test of sphericity**.



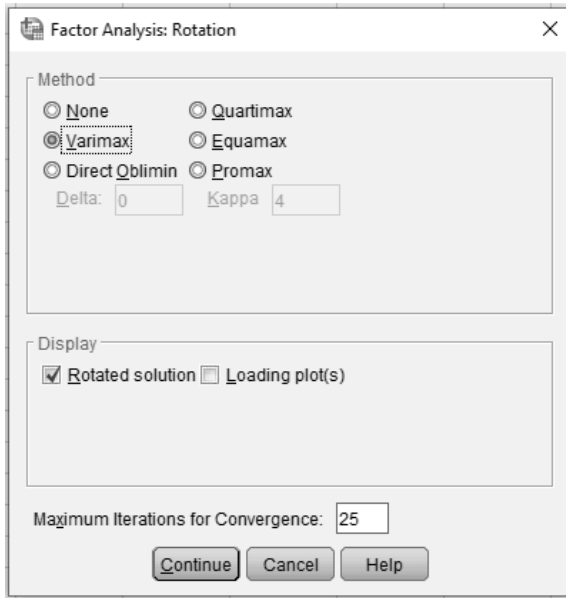
Gambar Penentuan factor analysis descriptive

Klik *extraction*. Pada bagian *Principal components*, centang *correlation matrix*, *Unrotated factor solution*, dan *Scree plot*. Adapun bagian *extract* pilih *Based on eigenvalue* seperti tersaji pada gambar



Gambar Penentuan factor analysis extraction

Pada bagian **Rotation**, pilih *varimax* dan *display* pilih *rotated solution*.



Gambar Penentuan factor analysis rotation

Setelah itu, klik *oke*. Berikut hasil analisis EFA yang siap diinterpretasikan.

1. Pengujian kelayakan analisis
Berikut hasil uji EFA.

Tabel 4.19 Uji KMO dan Bartlett

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.744
Bartlett's Test of Sphericity	Approx. Chi-Square	1176.345
	df	28
	Sig.	.000

Hasil EFA tentang kecukupan sampel menunjukkan nilai KHI-Kuadrat pada uji Bartlett sebesar 1176,345 dengan derajat kebebasan 28 dan nilai-p kurang dari 0,01. Artinya, ukuran sampel sebanyak 1261 yang digunakan telah cukup. Hal ini

dikuatkan dengan Kaiser Meyer Olkin *measure of sampling adequacy* (KMO) sebesar 0,744 yang lebih besar dari 0,5.

2. Menyajikan matrik korelasi

Perhatikan matrik korelasi pada tabel berikut ini.

Tabel 4.20 Anti-image Matrices

Anti-image Correlation	item1	.774 ^a	-.023	-.024	-.110	-.177	-.117	-.072	.011
	item2	-.023	.751 ^a	-.205	-.124	-.069	-.041	-.003	-.041
	item3	-.024	-.205	.788 ^a	-.058	-.015	-.172	-.107	-.082
	item4	-.110	-.124	-.058	.798 ^a	-.027	-.094	.003	-.077
	item5	-.177	-.069	-.015	-.027	.779 ^a	-.135	-.074	.003
	item6	-.117	-.041	-.172	-.094	-.135	.769 ^a	-.279	-.027
	item7	-.072	-.003	-.107	.003	-.074	-.279	.694 ^a	-.389
	item8	.011	-.041	-.082	-.077	.003	-.027	-.389	.697 ^a
a. Measures of Sampling Adequacy(MSA)									

Analisis dilanjutkan untuk mengetahui faktor-faktor pembentuk kejujuran. Anti-image Correaltion (AIC) menunjukkan variabel-variabel yang layak digunakan dalam melakukan analisis faktor. Hal ini dapat dilihat dari nilai Measure of Sampling Adequacy (MSA) dengan kode “a”. Hair (2010) menyatakan bahwa nilai MSA di atas 0,50 menunjukkan item layak untuk digunakan analisis faktor.

3. Melakukan ekstraksi

Ekstraksi faktor dilakukan dengan metode *principal component analysis* (PCA) dilanjutkan dengan rotasi faktor menggunakan metode varimax. Hasilnya terdapat pada Tabel 4.21.

Tabel 4.21 Communalitas

	Initial	Extraction
item1	1.000	.590
item2	1.000	.658
item3	1.000	.488
item4	1.000	.388
item5	1.000	.554
item6	1.000	.492
item7	1.000	.710
item8	1.000	.643
Extraction Method: Principal Component Analysis.		

Tabel 4.21 menunjukkan bahwa menunjukkan nilai komunalitas pada item3, item4, dan item6 kurang dari 0,50. Ketentuan item dipertahankan jika memiliki nilai *communalitas* lebih besar dari 0,5. Oleh karena itu dilakukan spesifikasi ulang analisis faktor dapat dengan mengambil salah satu mencakup opsi berikut, yaitu (1) menghapus variabel, (2) mengubah metode rotasi, atau (3) menambah atau mengurangi jumlah faktor (Hair, et. al. 2010). Pada pengolahan selanjutnya, dipilihlah obtion berupa mengeluarkan item dengan nilai komunalitas di bawah 0,50. EFA kemudian dilakukan beberapa kali sampai semua prasyarat menjadi signifikan. Hasilnya adalah sebagai berikut.

Tabel 4.22 KMO dan uji Bartlett (Tahap 2)

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.667
Bartlett's Test of Sphericity	Approx. Chi-Square	764.024
	df	10
	Sig.	.000

Tabel 4.22 Menunjukkan bahwa nilai KMO terpenuhi, yaitu 0,667 (lebih besar dari 0,5). Selain itu, nilai Khi-Kuadrat pada uji Bartlett sebesar 764,024 dengan derajat kebebasan 10 dan nilai-p kurang dari 0,01. Artinya, ukuran sampel telah cukup.

Tabel 4.23 Anti-image Matrices (tahap 2)

		item1	item5	item6	item8	item7
Anti-image Correlation	item1	.725 ^a	-.186	-.142	-.004	-.077
	item5	-.186	.724 ^a	-.152	-.006	-.078
	item6	-.142	-.152	.706 ^a	-.059	-.310
	item8	-.004	-.006	-.059	.635 ^a	-.407
	item7	-.077	-.078	-.310	-.407	.626 ^a
a. Measures of Sampling Adequacy(MSA)						

Anti-image Correlation (AIC) pada Tabel 4.23 menunjukkan bahwa semua item layak digunakan dalam melakukan analisis faktor. Hal ini dapat dilihat dari nilai *Measure of Sampling*

Adequacy (MSA) dengan kode “a”. di atas 0,50 yang menunjukkan item layak untuk digunakan analisis faktor. Langkah lanjutan yaitu dilakukannya *extraksi* menggunakan model verimax dengan hasil seperti pada Tabel 4.24

Tabel 4.24 Communalities (tahap 2)

	Initial	Extraction
item1	1.000	.574
item5	1.000	.570
item6	1.000	.505
item8	1.000	.694
item7	1.000	.710
Extraction Method: Principal Component Analysis.		

Tabel 4.24 menunjukkan bahwa menunjukkan nilai komunalitas semua item besar dari 0,5. Oleh karena itu, semua item memenuhi prasyarat dilakukannya analisis lanjut.

Tabel 4.25 Total Variance Explained

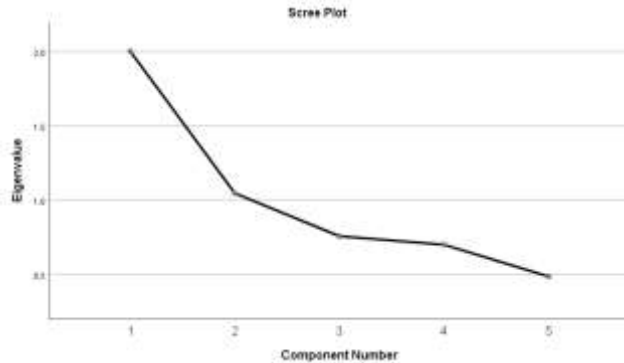
Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	2.005	40.098	40.098	2.005	40.098	40.098
2	1.048	20.957	61.055	1.048	20.957	61.055
3	.759	15.174	76.229			
4	.701	14.025	90.254			
5	.487	9.746	100.000			

Tabel 4.25 menjelaskan 3 hal, yaitu:

- a. Butir disusun berdasarkan urutan besarnya sumbangan terhadap keseluruhan kualitas butir.
- b. Dari ke-5 butir yang ada pada instrumen untuk mengukur kejujuran dapat diringkas menjadi 2 faktor karena beberapa butir mengukur dimensi yang sama. Marascuilo dan Cavin (1983) dalam Purwanto (2010) menyatakan bahwa faktor yang dipertahankan adalah yang memiliki eigenvalues di atas 1,00.

- c. Kedua faktor memberikan sumbangan total sebesar 61,055% dari keseluruhan kualitas instrument.

Hasil analisis yang tertuang pada Tabel.... Dipertegas oleh screen plot yang ada pada Gambar.... Menunjukkan bahwa grafik mengalami tingkat kecuraman yang besar pada 2 daerah.



Gambar Screen plot

4. Melakukan rotasi

Hasil analisis faktor analisis menggunakan rotasi terlihat pada Tabel Berdasarkan Tabel... maka dapat dilihat pengelompokan butir ke dalam faktor dan besarnya sumbangan seperti terlihat pada Tabel 4.26

Tabel 4.26 Rotated Component Matrix^a

	Component	
	1	2
item1	.064	.755
item5	.080	.751
item6	.544	.457
item8	.830	-.067
item7	.821	.189
Extraction Method: Principal Component Analysis.		
Rotation Method: Varimax with Kaiser Normalization.		
a. Rotation converged in 3 iterations.		

Tabel 4.26 pengelompokan butir ke dalam faktor dan besarnya sumbangan

Faktor	Butir	Muatan faktor
1	6	0,544
	8	0,830
	7	0,821
2	1	0,755
	5	0,751

5. Memberi nama faktor

Hasil rotasi menunjukkan pengelompokan dan besar sumbangan butir ke dalam faktor. Faktor-faktor masih berupa angka. Oleh karena itu, masing-masing faktor diberi nama sesuai dengan kesamaan ciri yang dimiliki oleh butir-butir yang mendukungnya.

Tabel 4.27 Pemberian nama faktor

Faktor	Butir	Nama/Label
1	6, 7, 8	Mempertahankan kepercayaan orang lain
2	1, 5	Tidak berbohong

Dengan demikian, maka dapat disimpulkan bahwa variabel kejujuran dibentuk dari dua faktor yaitu tidak berbohong dan mempertahankan kepercayaan orang lain dengan tidak berbuat curang.

Instrumen yang sudah memenuhi validitas isi dan validitas konstruk, siap digunakan untuk melakukan ujicoba lapangan menuju uji validitas empiris dan reliabilitas yang akan di bahas di BAB V.

BAB V

ANALISIS HASIL UJICоба

INSTRUMEN ANGKET

A. Prosedur Ujicoba Instrumen Angket

Instrumen yang akan digunakan dalam mengumpulkan data harus memenuhi uji validitas dan reliabilitas. Untuk memenuhinya maka perlu dilakukan ujicoba instrumen untuk kemudian menguji validitas dan reliabilitas.

Ujicoba instrumen angket sendiri merupakan langkah lanjutan dari pengembangan instrumen setelah instrumen dinyatakan memenuhi validitas isi oleh ahli dan dirakit kembali menjadi instrumen yang siap diujicobakan. Langkah ini ditempuh dengan cara meminta responden untuk mengisi angket. Responden yang mengisinya tentunya harus sesuai dengan subjek penelitian. Misalnya, angket akan digunakan untuk siswa kelas X di tingkat Sekolah Menengah Atas (SMA/MA), maka ujicoba angket dapat dilakukan terhadap siswa kelas X di SMA X, yaitu sekolah yang tidak akan digunakan dalam pengambilan data penelitian.

B. Analisis Hasil Ujicoba Instrumen Angket

Langkah awal yang perlu dilakukan peneliti saat data terkumpul adalah melakukan screening data. *Screening* dilakukan dengan tujuan melakukan drop terhadap jawaban responden yang tidak memenuhi aspek untuk diikutsertakan dalam analisis lanjutan. Berikut beberapa aspek tidak diikutsertakannya data dalam analisis data.

1. Identitas responden tidak ada (opsional tergantung variabel yang akan dituju dalam penelitian)
2. Tidak semua butir angket dijawab responden
3. Jawaban responden mengarah pada satu jawaban (skala)

4. Responden mengisi angket dalam waktu yang sangat singkat (tidak logis dengan alokasi waktu yang telah ditentukan)
5. Responden tidak memberikan jawaban yang konsisten (pernyataan positif dan negatif dijawab sama oleh responden).

Setelah proses screening, data diinput untuk kemudian dilanjutkan proses analisis data. Screening angket juga dapat dilakukan terhadap data yang telah terinput di program excel. Angket yang tidak lolos screening untuk selanjutnya didrop atau tidak diikutsertakan pada analisis data selanjutnya. Langkah lanjutannya yaitu merapikan data.

Data yang sudah rapi perlu dianalisis kembali dari sisi item angket. Jika peneliti menggunakan pertanyaan positif dan negatif negatif, maka peneliti perlu melakukan penyesuaian skala jawaban perntataan negatif yang diberikan responden.

Misalnya, peneliti menggunakan skala Likert dengan 4 skala. Pada pernyataan positif, skala yang digunakan 4 = selalu, 3 = sering, 2 = pernah, dan 1 = tidak pernah. Jawaban responden dari pernyataan positif dipertahankan seperti semula. Adapun jawaban responden terhadap pernyataan negatif harus dibalik. Agar lebih jelas, perhatikan tabel berikut.

Tabel 5.1 Konversi skoring jawaban responden pada pernyataan negatif

Kriteria	Jawaban responden pada pernyataan negatif	Data final responden
Selalu	4	1
Sering	3	2
Pernah	2	3
Tidak pernah	1	4

Analisis data lanjutan yaitu melakukan uji validasi empiris dan reliabilitas terhadap instrumen angket.

Contoh 5.1

Kegiatan screening

Berikut merupakan angket kejujuran yang sudah diisi responden. Angket terdiri atas beberapa butir positif dan negatif yang tersebar secara acak dengan kisi-kisi sebagai berikut.

Tabel 5.2 Indikator item angket kejujuran

No	Indikator	Indikator pernyataan	Nomor item	
			Positif	Negatif
1	Mengatakan sesuatu apa adanya (tidak berbohong)	Berbicara berdasarkan fakta di lapangan	1	-
		Tidak berbicara berdasarkan fakta	-	4
2	Tidak curang	Orang lain mengerjakan tugas	-	2
		Mengerjakan tugas sendiri	5	-
3	Mampu menjaga kepercayaan orang lain	Mengembalikan barang orang lain dalam keadaan baik dan bersih	6	-
		Tidak mengembalikan barang orang lain	-	3
		Meminta izin sebelum menggunakan barang orang lain	7	-
		Tidak meminta izin saat menggunakan barang orang lain	-	8

Item negatif terdiri dari item dengan nomor 2, 3, 4, dan 8. Jawaban responden terhadap butir angket negatif didasarkan pada skoring, yaitu 4 = sering, 3 = selalu, 2 = pernah, dan 1 = jarang. Skoring jawaban tersebut perlu dibalik terlebih dahulu menjadi yaitu 1

= sering, 2 = selalu, 3 = pernah, dan 4 = jarang seperti terlihat pada tabel berikut.

Tabel 5.3 Angket yang sudah diisi responden

Nomor	Item soal											
	Data asli								Data perubahan pada item			
	2	3	4	8	5	6	7	1	2	3	4	8
1	1	1	1	2	2	4	3	3	4	4	4	3
2	1	2	2	1	4	3	4	3	4	3	3	4
3	2	1	2	1	3	3	4	3	3	4	3	4
4	1	1	1	1	3	3	4	3	4	4	4	4
5	2	1	1	2	3	4	4	3	3	4	4	3
6	1	1	2	1	3	3	3	3	4	4	3	4
7	1	2	1	2	3	3	3	3	4	3	4	3
8	2	1	1	2	4	2	2	3	3	4	4	3
9	1	2	3	2	4	4	4	1	4	3	2	3
10	1	1	1	1	4	4	4	3	4	4	4	4
Dst	2	1	1	1	3	4	4	3	3	4	4	4

Data yang sudah diubah siap digunakan untuk melakukan validasi empiris dan reliabilitas butir soal.

1. Validitas empiris

Validitas empiris atau validitas kriteria adalah validitas yang ditentukan berdasarkan kriteria, yaitu kriteria internal maupun eksternal. Kriteria internal adalah instrumen itu sendiri yang menjadi kriteria, sedangkan kriteria eksternal merupakan hasil ukur instrumen di luar instrumen itu sendiri yang menjadi kriteria. Ukuran lain yang sudah dianggap baku atau dapat dipercaya dapat pula dijadikan sebagai kriteria eksternal (Sudaryono, Margono, dan Rahayu, 2013). Djaali dan Muljono (2004) mengartikan validitas empiris atau validitas konstruk sebagai validitas yang ditentukan berdasarkan data hasil ukur

instrumen yang bersangkutan baik melalui ujicoba maupun pengukuran yang sesungguhnya.

Suatu instrumen dikatakan memiliki validitas empiris atau belum jika sudah dilakukan penelusuran dari dua segi, yaitu daya ketepatan meramalnya (*predictive validity*) dan daya ketepatan bandingannya (*concurrent validity*).

Validitas ramalan suatu instrumen adalah suatu kondisi yang menunjukkan seberapa jauh sebuah instrumen telah dapat dengan tepat menunjukkan kemampuannya untuk meramalkan apa yang akan terjadi pada masa mendatang. Sudijono (2003) menyatakan bahwa untuk mengetahui apakah sebuah instrumen sudah memenuhi validitas ramalan atau belum dapat ditempuh dengan cara: mencari korelasi antara variabel yang sedang diuji validitas ramalannya dengan kriterium yang ada. Jika di antara kedua variabel terdapat korelasi positif yang signifikan, maka variabel yang sedang diuji validitas ramalannya dinyatakan memiliki daya ramal yang tepat. Artinya apa yang telah diramalkan, betul-betul telah terjadi secara nyata dalam praktik.

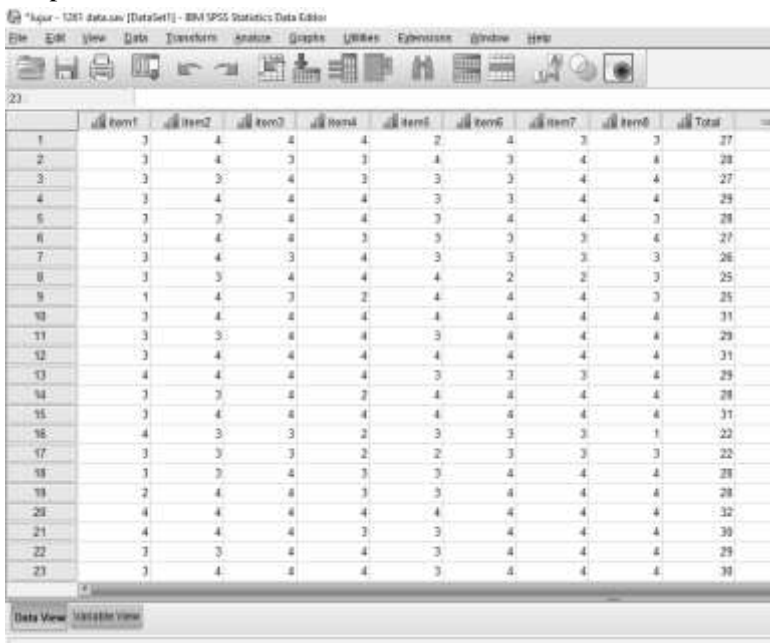
Instrumen dapat dikatakan telah memiliki validitas bandingan jika instrumen tersebut dalam kurun waktu yang sama dengan secara tepat telah mampu menunjukkan adanya hubungan yang searah antara penilaian awal dan berikutnya. Sudijono (2003) menunjukkan cara pengujian validitas bandingan dengan cara membandingkan data yang mencerminkan pengalaman masa lalu dan hasil penilaian sekarang. Jika hasil penilaian memiliki hubungan yang searah maka instrumen tersebut dapat dikatakan telah memiliki validitas bandingan.

Berdasarkan pembahasan di atas, dapat disimpulkan bahwa validitas ramalan dan validitas bandingan merupakan validitas yang ditinjau hubungannya dengan validitas kriterium. Jika kriterium yang dihubungkan terdapat pada waktu yang akan datang, maka masuk dalam validitas ramalan. Sebaliknya, jika kriteriumnya terdapat pada saat sekarang atau pada kurun waktu yang bersamaan dengan pengukur yang sedang diuji validitasnya maka disebut validitas bandingan.

Contoh 5.2.

Perhitungan validitas empiris angket living values (kejujuran) dalam proses pembelajaran yang telah diisi oleh 1261 responden. Jumlah item angket sebanyak 8. Penentuan validitas empiris menggunakan korelasi Pearson

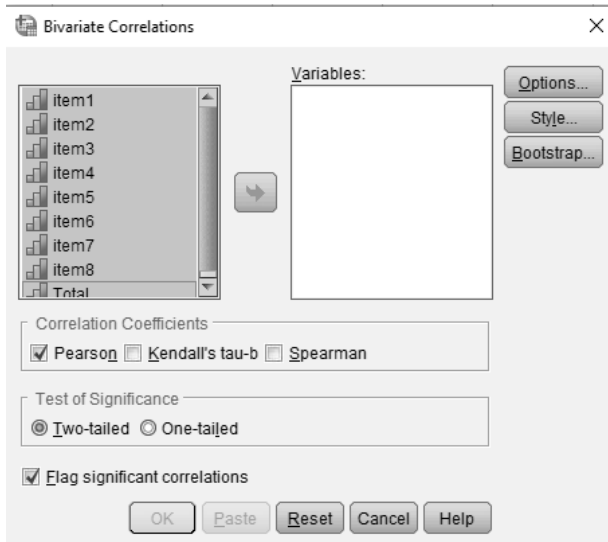
Setelah data berupa 8 item angket yang telah diisi oleh 1261 responden diinput ke SPSS pada data view. Langkah selanjutnya yaitu mencari validitas butir angket melalui korelasi Pearson dengan sebelumnya melengkapi data dengan skor total masing-masing responden dalam menjawab angket seperti terlihat pada tampilan berikut.



The screenshot shows the SPSS Data Editor interface. The title bar reads "Kajur - 1261 data.sav [DataSet1] - IBM SPSS Statistics Data Editor". The menu bar includes File, Edit, View, Data, Transform, Analyze, Graphs, Utilities, Extensions, Window, and Help. The toolbar contains various icons for file operations and analysis. The main data grid has 23 rows and 9 columns. The columns are labeled "item1", "item2", "item3", "item4", "item5", "item6", "item7", "item8", and "Total". The data values are integers ranging from 1 to 5. The "Total" column represents the sum of the 8 items for each respondent.

	item1	item2	item3	item4	item5	item6	item7	item8	Total
1	3	4	4	4	2	4	3	3	27
2	3	4	3	3	4	3	4	4	28
3	3	3	4	3	3	3	4	4	27
4	3	4	4	4	3	3	4	4	29
5	3	3	4	4	3	4	4	3	28
6	3	4	4	3	3	3	3	4	27
7	3	4	3	4	3	3	3	3	26
8	3	3	4	4	4	2	2	3	25
9	1	4	3	2	4	4	4	3	25
10	3	4	4	4	4	4	4	4	31
11	3	3	4	4	3	4	4	4	29
12	3	4	4	4	4	4	4	4	31
13	4	4	4	4	3	3	3	4	29
14	3	3	4	2	4	4	4	4	28
15	3	4	4	4	4	4	4	4	31
16	4	3	3	2	3	3	3	1	22
17	3	3	3	2	2	3	3	3	22
18	1	3	4	3	3	4	4	4	26
19	2	4	4	3	3	4	4	4	28
20	4	4	4	4	4	4	4	4	32
21	4	4	4	3	3	4	4	4	30
22	3	3	4	4	3	4	4	4	29
23	3	4	4	4	3	4	4	4	30

Langkah selanjutnya yaitu klik **analyze** → **correlate** → **Bivariate**



Masukkan semua item dan total ke kolom *Variables*. Pada *correlation coefficients* klik *Pearson*. Pada kolom *test of Significance*, pilih *Two-tailed*. Selanjutnya beri tanda centang pada *Flag significant correlations*. Kemudian, klik **OK**. Berikut hasilnya.

Tabel 5.4 Correlation angket kejujuran

		item1	item2	item3	item4	item5	item6	item7	item8	Total
item1	Pearson Correlation	1	.104**	.131**	.172**	.242**	.235**	.190**	.101**	.496**
	Sig. (2-tailed)		.000	.000	.000	.000	.000	.000	.000	.000
	N	1261	1261	1261	1261	1261	1261	1261	1261	1261
item2	Pearson Correlation	.104**	1	.269**	.185**	.132**	.168**	.136**	.132**	.456**
	Sig. (2-tailed)	.000		.000	.000	.000	.000	.000	.000	.000
	N	1261	1261	1261	1261	1261	1261	1261	1261	1261
item3	Pearson Correlation	.131**	.269**	1	.162**	.129**	.304**	.275**	.222**	.541**
	Sig. (2-tailed)	.000	.000		.000	.000	.000	.000	.000	.000
	N	1261	1261	1261	1261	1261	1261	1261	1261	1261
item4	Pearson Correlation	.172**	.185**	.162**	1	.113**	.196**	.139**	.152**	.489**
	Sig. (2-tailed)	.000	.000	.000		.000	.000	.000	.000	.000
	N	1261	1261	1261	1261	1261	1261	1261	1261	1261

item5	Pearson Correlation	.242**	.132**	.129**	.113**	1	.245**	.195**	.106**	.510**
	Sig. (2-tailed)	.000	.000	.000	.000		.000	.000	.000	.000
	N	1261	1261	1261	1261	1261	1261	1261	1261	1261
item6	Pearson Correlation	.235**	.168**	.304**	.196**	.245**	1	.413**	.241**	.635**
	Sig. (2-tailed)	.000	.000	.000	.000	.000		.000	.000	.000
	N	1261	1261	1261	1261	1261	1261	1261	1261	1261
item7	Pearson Correlation	.190**	.136**	.275**	.139**	.195**	.413**	1	.462**	.623**
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000		.000	.000
	N	1261	1261	1261	1261	1261	1261	1261	1261	1261
item8	Pearson Correlation	.101**	.132**	.222**	.152**	.106**	.241**	.462**	1	.537**
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000	.000		.000
	N	1261	1261	1261	1261	1261	1261	1261	1261	1261
Total	Pearson Correlation	.496**	.456**	.541**	.489**	.510**	.635**	.623**	.537**	1
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000	.000	.000	
	N	1261	1261	1261	1261	1261	1261	1261	1261	1261

** . Correlation is significant at the 0.01 level (2-tailed).

Data pada Tabel 5.4 kemudian dianalisis untuk mengetahui validitas tiap item dengan cara membandingkan koefisien Pearson correlation setiap item angket dengan R tabel pada taraf signifikansi 5%. Jika koefisien Pearson correlation hasil perhitungan lebih besar dibanding R tabel, maka item dinyatakan valid. Cara lainnya dengan melihat nilai signifikansi. Jika signifikansi item kurang dari 0,05, maka item dinyatakan valid. Berikut tabel *r correlation produk moment*.

Tabel 5.5 Nilai *r correlation produk moment.*

Sumber: Sugiono, 1999

N	Taraf Signif		N	Taraf Signif		N	Taraf Signif	
	5%	1%		5%	1%		5%	1%
3	0,997	0,999	27	0,381	0,467	55	0,266	0,345
4	0,950	0,990	28	0,374	0,478	60	0,254	0,330
5	0,878	0,959	29	0,387	0,470	65	0,244	0,317
6	0,811	0,917	30	0,361	0,483	70	0,235	0,306
7	0,754	0,874	31	0,355	0,458	75	0,227	0,296
8	0,707	0,834	32	0,349	0,449	80	0,220	0,288
9	0,668	0,798	33	0,344	0,442	85	0,213	0,278
10	0,632	0,765	34	0,339	0,436	90	0,207	0,270
11	0,602	0,735	35	0,334	0,430	95	0,202	0,263
12	0,576	0,708	36	0,329	0,424	100	0,195	0,258
13	0,553	0,684	37	0,325	0,418	125	0,176	0,230
14	0,532	0,661	38	0,320	0,413	150	0,159	0,210
15	0,514	0,641	39	0,316	0,408	175	0,148	0,194
16	0,497	0,623	40	0,312	0,403	200	0,138	0,181
17	0,482	0,606	41	0,308	0,398	300	0,113	0,148
18	0,468	0,590	42	0,304	0,393	400	0,098	0,128
19	0,456	0,575	43	0,301	0,389	500	0,088	0,115
20	0,444	0,561	44	0,297	0,384	600	0,080	0,105
21	0,433	0,549	45	0,294	0,380	700	0,074	0,097
22	0,423	0,537	46	0,291	0,376	800	0,070	0,091
23	0,413	0,526	47	0,288	0,372	900	0,065	0,086
24	0,404	0,515	48	0,284	0,368	1000	0,062	0,081
25	0,396	0,505	49	0,281	0,364			
26	0,388	0,496	50	0,279	0,361			

Berdasarkan Tabel 5.5, diketahui bahwa pada $N = 1000$ nilai koefisien r pada taraf signifikansi 5% adalah 0,062 dan 1% sebesar 0,081. Nilai koefisien pada taraf signifikansi 5% yang nantinya dibandingkan dengan nilai koefisien r hitung.

Tabel 5.6 Penentuan validasi empiris item angket kejujuran berdasarkan korelasi Pearson

Item	Pearson correlation	R Tabel	Signifikansi	Keterangan
1	0,496	0,062	0,000	Valid
2	0,456	0,062	0,000	Valid
3	0,541	0,062	0,000	Valid
4	0,489	0,062	0,000	Valid
5	0,510	0,062	0,000	Valid
6	0,635	0,062	0,000	Valid
7	0,623	0,062	0,000	Valid
8	0,537	0,062	0,000	Valid

Berdasarkan Tabel 5.6 diketahui nilai r hitung pada semua item lebih besar dari r tabel. Selain itu, nilai signifikansi semua item kurang dari 0,05. Oleh karena itu, dapat dikatakan bahwa semua item memenuhi validitas empiris. Oleh karena itu, semua item angket siap diuji reliabilitas untuk mengetahui sejauhmana hasil pengukuran dapat dipercaya.

2. Reliabilitas Angket

Reliabilitas berasal dari kata *reliability* berarti sejauhmana hasil suatu pengukuran dapat dipercaya. Suatu hasil pengukuran hanya dapat dipercaya jika dalam beberapa kali pelaksanaan pengukuran terhadap kelompok subjek yang sama, diperoleh hasil pengukuran yang relatif sama, selama aspek yang diukur dalam diri subjek memang belum berubah. (Sudaryono, Margono, dan Rahayu (2013). Hal yang sama juga disampaikan oleh Mardapi (2017) dan Purwanto (2010) bahwa reliabilitas diartikan sebagai keterpercayaan yang berhubungan dengan tetapan dan konsistensi. Sebuah instrument dikatakan dapat dipercaya atau reliabel jika memberikan hasil pengukuran yang relatif konsisten.

Retnawati (2017) menambahkan bahwa reliabilitas juga terkait dengan kesalahan pengukuran. Reliabilitas tinggi

menunjukkan kesalahan yang kecil dalam memperoleh hasil pengukuran. Semakin besar reliabilitas suatu instrumen maka semakin kecil kesalahan pengukuran. Sebaliknya, semakin kecil reliabilitas skor, akan semakin besar kesalahan pengukuran. Kesalahan pengukuran itu sendiri dapat disebabkan oleh beberapa faktor, di antaranya: (1) karakteristik instrumen yang digunakan yang dapat disebabkan karena penyusunan dan pelaksanaan pengukuran yang tidak mengikuti aturan baku, (2) kualitas butir dalam instrumen tidak baik, (3) adanya kerjasama responden selama mengisi instrumen, (4) butir-butir instrumen yang meragukan, (5) keadaan peserta selama merespon instrumen, seperti responden mengalami kelelahan fisik maupun psikis, memiliki problem pribadi, memiliki motivasi kurang, (6) lingkungan tempat penyelenggaraan pengukuran yang kurang mendukung atau kombinasi dari segala permasalahan tersebut.

Reliabilitas instrumen diekspresikan secara numeric dalam bentuk koefisien yang besarnya -1,00 hingga 1,00. Koefisien tinggi menunjukkan reliabilitas tinggi, sebaliknya koefisien rendah menunjukkan tingkat reliabilitas instrumen rendah.

Kerlinger (1996) menyampaikan beberapa batasan terkait dengan reliabilitas. (1) Reliabilitas dicapai apabila kita mengukur himpunan objek yang sama berulang kali dengan instrumen yang sama atau serupa dan memberikan hasil yang sama atau serupa. (2) Reliabilitas dicapai apabila ukuran yang diperoleh dari suatu instrumen pengukur adalah ukuran “yang sebenarnya” untuk sifat yang diukur. (3) Reliabilitas dicapai dengan meminimalkan alat pengukuran yang terdapat dalam suatu instrumen pengukur. Instrumen yang tidak reliable tidak dapat digunakan untuk mengumpulkan data karena tidak memberikan informasi apapun.

Purwanto (2010) mengelompokkan reliabilitas menjadi 2 kelompok, yaitu (1) kelompok pertama yang melakukan pengukuran ketika instrumen diuji beberapa kali dan (2) kelompok kedua yang menyatakan reliabilitas sebagai konsistensi internal hasil pengukuran butir-butir instrumen yang dapat diukur

hanya satu kali. Berikut penjelasan lebih detail masing-masing kelompok.

a. Kelompok pertama

Kelompok pertama mendefinisikan reliabilitas sebagai kestabilan hasil pengukuran ketika instrumen diuji beberapa kali. Reliabilitas yang masuk dalam golongan ini adalah metode tes-retest dan metode paralel.

1) Metode Tes ulang (*test-retest method*)

Metode tes ulang yaitu metode pengujian reliabilitas yang dilakukan dengan menyajikan sebuah perangkat instrumen kepada kelompok responden sebanyak dua kali. Hasil pengukuran kedua pengujian selanjutnya dikorelasikan. Sebuah instrumen dikatakan reliabel jika dua kali pengujian menunjukkan hasil yang stabil. Stabilitas ditunjukkan oleh korelasi antara skor yang diperoleh dari kedua pengujian (Purwanto, 2010). Retnowati (2017) menyatakan bahwa metode ini akan dapat sesuai dengan tujuannya jika keadaan sumber yang diukur tetap dan tidak mengalami perubahan pada saat pengukuran pertama dan kedua. Retnowati (2017) lebih menyarankan beberapa hal yang harus diperhatikan dalam mengestimasi koefisien dengan teknik tes-retest ini. Pertama yaitu jangka waktu pelaksanaan kedua pengukuran perlu dipertimbangkan dengan baik. Jarak tes yang terlalu dekat dapat mengakibatkan terjadinya *testing effect*, yaitu responden masih teringat dengan pengukuran pertama. Sebaliknya, jika terlalu lama, korelasi kedua skor akan menjadi semakin rendah. Faktor kedua adalah adanya stabilitas kinerja yang diukur dengan instrumen tersebut. Sebagai saran, jarak kedua pengukuran sebaiknya tidak terlalu jauh, misalnya tidak sampai satu bulan.

2) Metode paralel (*equivalent/alternate form*)

Metode paralel yaitu pengujian reliabilitas yang dilakukan dengan cara membuat dua perangkat instrumen yang paralel atau ekuivalen dan mengujikan sekaligus. Dua perangkat paralel yaitu dua perangkat instrumen yang dikembangkan dari spesifikasi yang sama meliputi: jumlah butir, pelaksanaan, bentuk, waktu, ujicoba, peserta ujicoba, dan kisi-kisi. Spesifikasi instrumen merupakan detail rancangan yang mengarahkan pada penulisan butir-butir instrumen yang akan digunakan untuk pengumpulan data. Perangkat yang dihasilkan akan paralel jika dikembangkan berdasarkan spesifikasi yang sama, terlepas siapa dan kapan pengembangan instrumen dilakukan. Selanjutnya, koefisien reliabilitas ditentukan dengan menghitung koefisien korelasi skor responden pada kedua instrumen (Purwanto, 2010).

Langkah-langkah pembuktian reliabilitas menggunakan model paralel menurut Retnawati (2017) yaitu: (1) menyiapkan dua instrumen yang paralel, (2) menentukan subjek untuk mengujicobakan instrumen, (3) melaksanakan pengukuran dengan mengadministrasikan instrumen, (4) melaksanakan penyekoran pada setiap jawaban responden terhadap kedua perangkat, (5) menghitung koefisien korelasi dari skor kedua perangkat. Semakin tinggi skor jawaban responden pada kedua instrumen yang digunakan menunjukkan bahwa reliabilitas paralel instrumen berada pada kategori baik.

b. Kelompok kedua

Kelompok kedua memaknai reliabilitas sebagai konsistensi internal hasil pengukuran butir-butir instrument. Metode ini terbagi menjadi 2. Jika jumlah butir instrumen genap maka metode dapat menggunakan metode belah dua. Sebaliknya, jika jumlah butir ganjil maka formula yang dapat digunakan antara lain Kuder-Richardson, Hoyt, dan Alpha Cronbach.

1) Metode belah dua

Metode belah dua yaitu uji reliabilitas yang dilakukan jika peneliti yang hanya mengembangkan satu perangkat instrumen dan tidak ingin mengujicobakan dua kali. Metode ini dilakukan dengan cara membagi butir perangkat menjadi dua belahan, selanjutnya mengkorelasikan skor total kedua belahan. Setiap butir dalam instrumen harus mengukur hal yang sama, sehingga korelasi antara mereka harus tinggi. Oleh karena butir akan dibagi menjadi dua yang sama besar, maka jumlah butir instrumen harus genap (Purwanto, 2010). Retnowati (2017) menyebutkan beberapa formula reliabilitas yang menggunakan metode belah dua antara lain: (a) rumus Spearman-Brown, (b) rumus Flagnan, dan (c) rumus Rulon dengan persamaan berikut.

Reliabilitas dengan rumus Spearman-Brown

$$r_i = \frac{2r_b}{1 + r_b}$$

$$r_b = \frac{N \sum XY - \sum X \sum Y}{\sqrt{\{N \sum X^2 - (\sum X)^2\} \{N \sum Y^2 - (\sum Y)^2\}}}$$

Keterangan:

r_i = koefisien reliabilitas skor instrumen,

r_b = koefisien korelasi antara dua belahan instrumen,

N = banyaknya responden,

X = belahan pertama,

Y = belahan kedua

Reliabilitas dengan rumus Flanagan

$$r_i = \frac{2\sigma_{12}}{1 + \sigma_X^2} = \frac{2r_{12}\sigma_1\sigma_2}{\sigma_1^2 + \sigma_2^2 + 2r_{12}\sigma_{12}}$$

Keterangan:

r_i = koefisien reliabilitas skor instrumen,

r_{12} = koefisien korelasi antara dua belahan instrumen,

σ_1^2 = varians belahan pertama,

σ_2^2 = varians belahan kedua,

σ_X^2 = varians skor total

Reliabilitas dengan rumus Rulon

$$r_i = 1 - \frac{\sigma_d^2}{\sigma_t^2}$$

Keterangan:

r_i = reliabilitas instrumen,

σ_t^2 = varians total atau varians skor total;

σ_d^2 = varians dari perbedaan skor kedua belahan (*varians difference*);

d = skor pada belahan awal dikurangi skor pada belahan akhir

Kuder-Richardson

Kuder-Richardson merupakan salah satu uji reliabilitas yang dilakukan jika jumlah butir instrumen berjumlah ganjil. Di antara beberapa rumus Kuder dan Richardson, KR-20 dan KR-21 merupakan rumus yang paling banyak digunakan.

Retnowati (2017) menyatakan bahwa KR-20 digunakan untuk analisis butir dikotomi, seperti 1-0, benar-salah, ya-tidak, hidup-mati. Adapun KR-21 digunakan untuk instrumen yang menggunakan skor 0 dan 1, skala Likert (musalnya, 1-2-3-4-5), dan penilaian tes dalam bentuk uraian.

Rumus KR-20

$$r_{11} = \left(\frac{n}{n-1} \right) \left(\frac{s_t^2 - \sum pq}{s_t^2} \right)$$

Keterangan:

n = jumlah butir,

s_t^2 = varians total,

p = proporsi skor yang diperoleh,

q = proporsi skor maksimum dikurangi skor yang diperoleh

Rumus KR-21

$$r_{11} = \left(\frac{n}{n-1} \right) \left(1 - \frac{M(n-M)}{ns_t^2} \right)$$

Keterangan:

n = jumlah butir,

M = rata-rata skor total,

s_t^2 = varians total

Hoyt

Hoyt merupakan salah satu metode perhitungan koefisien reliabilitas lain yang dilakukan jika jumlah butir instrumen berjumlah ganjil dengan persamaan berikut.

$$r_{11} = 1 - \frac{V(s)}{V(r)}$$

Keterangan:

r_{11} = koefisien reliabilitas,

$V(r)$ = varians responden,

$V(s)$ = varians sisa

Alpha Cronbach

Alpha Cronbach merupakan salah satu metode perhitungan koefisien reliabilitas lain yang dilakukan jika jumlah butir instrumen berjumlah ganjil dengan persamaan berikut.

$$r_{11} = \left(\frac{n}{n-1} \right) \left(1 - \frac{\sum s_i^2}{\sum s_t^2} \right)$$

Keterangan:

n = jumlah butir,

s_i^2 = varians butir,

s_t^2 = varians total

Retnowati (2017) menambahkan bahwa Alpha Cronbach dapat digunakan untuk mengestimasi reliabilitas instrumen yang skornya 1 dan 0, skala politomus misalnya skala Likert, atau soal dalam bentuk uraian dimana skor maksimum tergantung pada peneliti.

Penentuan reliable tidaknya suatu instrumen ditentukan oleh koefisien reliabilitas dari uji reliabilitas yang dilakukan. Purwanto (2010) menyatakan bahwa instrumen dapat dinyatakan reliable jika koefisien yang diperoleh melalui perhitungan menggunakan metode pengujian reliabilitas tertentu lebih besar dibandingkan dengan batas keputusan reliabilitas. Gronlund dan Linn (1990) menyatakan bahwa penentuan reliabilitas instrumen ditentukan oleh koefisien reliabilitas yang akan sangat bervariasi sesuai keputusan yang akan dibuat, meliputi: (1) tingkat pentingnya keputusan, (2) dapat tidaknya keputusan diperbaiki dalam waktu cepat, (3) jaminan yang kita butuhkan sehubungan dengan keputusan yang dibuat. Tingkat pentingnya keputusan berhubungan dengan seberapa penting konsekuensi bagi responden. Jika membutuhkan konsekuensi yang sangat penting bagi responden, maka instrumen dituntut memiliki koefisien reliabilitas yang sangat tinggi. Terkait dengan kecepatan memperbaiki keputusan, semakin lama waktu yang diperlukan untuk memperbaiki instrumen maka semakin tinggi koefisien reliabilitas

yang diperlukan. Adapun jaminan yang lebih besar mensyaratkan reliabilitas yang lebih tinggi.

Berbeda dengan pandangan ahli sebelumnya, Aiken (1995) menyatakan bahwa penentuan tinggi rendahnya koefisien reliabilitas suatu instrumen sangat tergantung pada apa yang akan dilakukan atas skor. Jika skor digunakan untuk menentukan apakah dua kelompok berbeda signifikan maka koefisien reliabilitas 0,65 sudah memberikan kontribusi dalam keputusan. Sebaliknya, jika skor digunakan untuk membandingkan penampilan individu yang berbeda maka koefisien reliabilitas minimal 0,85.

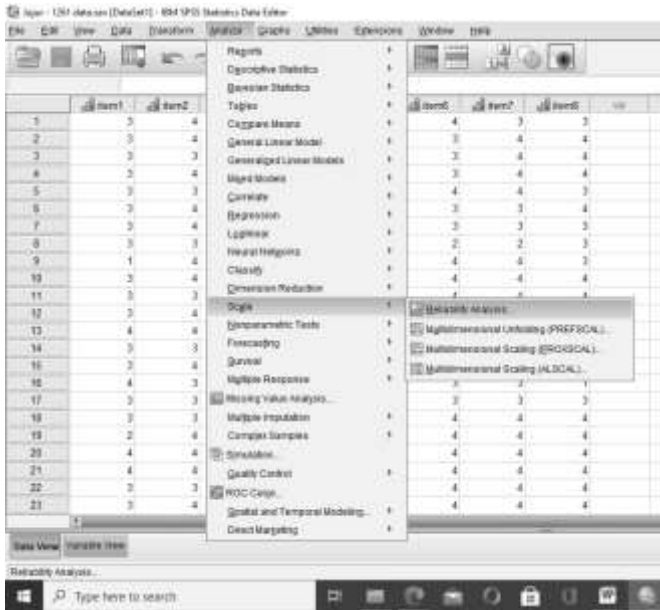
Santayasa (2014) memberikan ketentuan berbeda terkait reliabilitas instrumen berdasarkan nilai koefisien reliabilitasnya dengan membaginya menjadi 4 kriteria, yaitu 0,80 - 1,00 (sangat tinggi); 0,60 - 0,80 (tinggi); 0,40 - 0,60 (sedang); 0,20-0,40 (rendah).

Contoh 5.3

Perhitungan reliabilitas angket living values (kejujuran) dalam proses pembelajaran menggunakan Alpha Cronbach (Sarah, Prasetyo, Wilujeng, 2019)

Reliabilitas angket kejujuran sebanyak 8 item ditentukan melalui analisis Alpha Cronbach setelah angket diujicobakan ke 1261 responden dan dinyatakan semuanya memenuhi validitas empiris. Penentuan reliabilitas menggunakan Alpha Cronbach dilatarbelakangi jenis data yang digunakan yaitu data angket dengan 4 skala pada skala Likert. Penentuan koefisien Alpha Cronbach dilakukan menggunakan bantuan SPSS vers 25.00. Berikut langkahnya.

Data dimasukkan ke SPSS, Klik **analyze** → **Scale** → **Reliability Analyze**



Gambar



Gambar

Masukkan semua data ke kolom item. Pastikan pada bagian **Model** tertera **Alpha**. Setelah itu, klik **OK**. Berikut hasilnya.

Tabel. Hasil uji reliabilitas angket kejujuran menggunakan Alpha Cronbach

Reliability Statistics	
Cronbach's Alpha	N of Items
.660	8

Berdasarkan Santyasa (2014), dapat disimpulkan bahwa angket kejujuran yang terdiri dari 8 item memenuhi kriteria reliable dengan kriteria tinggi.

Penerapan Isu Lingkungan 5.1

Penggunaan formula Alpha Cronbach juga terdapat dalam penelitian berjudul “Pengembangan Instrumen Self Assessment Technological Pedagogical Content Knowledge (TPACK) Calon Guru Matematika Berbasis HOTS” yang dilakukan oleh Ahmad Zaeni , Wardani Rahayu, dan Makmuri yang diterbitkan di Teorema: Teori dan Riset Matematika, 5(2), 59–68, Maret 2021. Bahasan secara lengkap dapat dilihat langsung melalui link.

<https://jurnal.unigal.ac.id/index.php/teorema/article/view/4960/3857>

DAFTAR PUSTAKA

- Aiken, L. R. (1985). Three coefficients for analyzing the reliability and validity of ratings. *Educational and Psychological Measurement*, 45: 131-142.
- Ali, Muhammad. (2009). *Metode Penelitian Pendidikan*. Bandung: Angkasa.
- Allen, M. J. & Yen, M. W. (1979). *Introduction to measurement theory*. Monterey, CA: Books/Cole Publishing Company.
- Andersen, L. W. (1981). *Assessing affective characteristic in the school*. Boston: Allyn and Bacon.
- Andriani, N & Rasto. (2019). Motivasi belajar sebagai determinan hasil belajar siswa. *Jurnal Pendidikan Manajemen Perkantoran*, 4(1): 80-86.
- Anugrahana, A. (2020). Hambatan, Solusi dan Harapan: Pembelajaran Daring Selama Masa Pandemi Covid-19 Oleh Guru Sekolah Dasar. *Scholaria: Jurnal Pendidikan dan Kebudayaan*, 10(3): 282-289
- Ardiyanto, A., Mulyadin, T., Santi, A. M., & Dharma, I. G. (2021). Online Classrooms During COVID-19 Pandemic: A Survey of Industrial Engineering Instructors in Indonesia. *Higher Learning Research Communications*, 11 (1). DOI:[10.18870/hlrc.v11i1.1232](https://doi.org/10.18870/hlrc.v11i1.1232)
- Bunn, D. N., Caudill, S. B., & Gropper, D. M. (1992) Crime in the Classroom: An Economic Analysis of Undergraduate Student Cheating Behavior. *The Journal of Economic Education*, 23/3, 197-207.
- Cresswell, J.W. (2012). *Educational Research: Planning, Conductioning, and Evaluating Quantitative and Qualitative Research* (4th ed). New York: Pearson Merrill Prentice Hall.

- Daskon, C. D. (2015). Significance of cultural values in securing asset portfolios of rural communities. *Sri Lanka Journal of Social Sciences*, 38/1, 31-50.
- DePaulo, B.M.; Kashy, D. A.; Kirkendol, S. E.; Wyer, M. M.; & Epstein, J. A. (1996). Lying in Everyday Life. *Journal of Personality and Social Psychology*, 70(5): 979-995.
- Djaali, P. M. (2004). *Pengukuran dalam Bidang Pendidikan*. Jakarta: Program Pascasarjana Universitas Negeri Jakarta.
- Ehrlich, T. & Fu, E. (2013). *Cheating in School and Colleges: What To Do About It*. Retrieved 3 October, 2018 from <https://www.forbes.com/sites/ehrllichfu/2013/08/22/cheating-in-schools-and-colleges/#43fa464837aa>.
- Fernandes, H. J. X. (1984). *Evaluation of education program*. Jakarta: National Education Planning, Evaluation, and Curriculum Development.
- Fleming, J. S. (1980). Cheating Behavior, Subject Variables, and Components of the Internal-External Scale under High and Low Risk Conditions. *The Journal of Educational Research*, 74/2, 83-87. <https://www.jstor.org/stable/27507244>
- Gächter, S. & Schulz, J. F. (2016). Intrinsic honesty and the prevalence of rule violations across societies. *Nature*, 531, 496–499.
- Gay, L.R., Milles, G. E., Airasian, P. W. (2012). *Educational Research: Competencies for Analysis and Applications* (10th ed). New Jersey: Pearson Education, Inc.
- Gregory, R. J. (2007). *Psychological testing: history, principles, and applications*: Boston: Pearson.
- Hadiati, S., Kuswanto, H., Rosana, D., & Pramuda, A. (2019). The Effect of Laboratory Work Style and Reasoning with Arduino to Improve Scientific Attitude. *International Journal of Instruction*, 12(2), 321-336. <https://doi.org/10.29333/iji.2019.12221a>

- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis* (7th ed). Prentice Hall [versi elektronik].
- Hidayati, K., Budiyono, & Sugiman. (2018). Development and Validation of Student's Responsibility Scale on Mathematics Learning Using Subject Scalling Model. *International Journal of Instruction*, 11/4, 499-512.
- Ismail, I. A. (2009). *Pilar-Pilar Takwa Doktrin, Pemikiran, Hikmat, dan Pencerahan Spiritual* [Pillars of Piety, Doctrine, Thought, Wisdom, and Spiritual Enlightenmen]. Jakarta: PT RajaGrafindo Persada.
- Kemendiknas. (2010). *Pengembangan Pendidikan dan Karakter Bangsa* [National Education and Character Development]. Jakarta: Pusat Kurikulum.
- Kerlinger, F.N. 1986. *Foundation of Behavioral Research*. Yogyakarta : Penerbit Fakultas Psikologi UGM.
- Kirschenbaum, H. (1995). *100 Ways to Enhance Values and Morality in Schools and Youth Setting*. Boston: Allyn and Bacon.
- Klieger, Aviva & Sherman, G. (2015). Physics textbooks: do they promote or inhibit students' creative thinking. *Physics Education* 50(3): 305-309.
- Koesoema, D. A. 2015. *Pendidikan karakter utuh dan menyeluruh*. Yogyakarta: Kanisius.
- Komalasari, K. (2012). The Living Values-Based Contextual Learning to Develop the Students' Character. *Journal of Social Sciences*, 8/2, 246-251.
- Komalasari, K., & Saripudin, D. (2018). The Influence of Living Values Education-Based Civic Education Textbook on Student's Character Formation. *International Journal of Instruction*, 11/1, 395-410.
- Komalasari, K., Saripudin, D., & Masyitoh, I. S. (2014). Living Values Education Model in Learning and Extracurricular

- Activities to Construct the Students' Character. *Journal of Education and Practice*, 5/7, 166-174.
- Krathwohl, D. R., Bloom, B.S., & Masia, B.B. (1964). *Taxonomy of educational objectives: Handbook II: Affective Domain*. New York: David Mc Kay.
- LaForgia, J. (1988). The Affective Domain Related to Science Education and Its Evaluation. *Science Education*, 72, 407-421.
- Lickona, T. (2013). *Pendidikan Karakter: Panduan Lengkap Mendidik Siswa Menjadi Pintar dan Baik* [Character Education: A Complete Guide to Educating Students to Be Smart and Good]. Bandung: Nusa Media.
- Linn, R. L. & Gronlund, N. E. (1995). *Measurement and assessment in teaching (7th ed)*. EnglewoodCliffs, NJ: Prentice-Hall.
- Madjid, N. (1992). *Islam: doktrin dan peradaban*. Jakarta: Paramadina.
- Mahmud. 2011. *Metode Penelitian Pendidikan*. Bandung: Pustaka Setia.
- Mardapi, D. (2008). *Teknik Penyusunan Tes dan Nontes* [Test and Non-Test Preparation Techniques]. Yogyakarta: Mitra Cendekia Press.
- _____. (2017). *Pengukuran, Penilaian dan Evaluasi Pendidikan* [Educational Measurement, Assessment and Evaluation]. Yogyakarta: Parama Publishing.
- Mazar, N., Amir, O., & Ariely, D. (2008). The dishonesty of Honest People: A Theory of Self-Concept Maintenance. *Journal of Marketing Research*. XLV, 633-644.
- Moten, A. R. (2014). Academic dishonesty and misconduct: Curbing plagiarism in the Muslim world. *Intellectual Discourse*, 22/2, 167-189.
- Mu'in, F. (2013). *Pendidikan karakter konstruksi teori & praktek*. Jogjakarta: Ar-Ruzz Media.

- Muhasim. (2017). Budaya Kejujuran dalam Menghadapi Perubahan Zaman: Studi Fenomenologi Masyarakat Islam Modern [Culture of Honesty in Facing the Changes of the Age: Phenomenology Study of Modern Islamic Society]. *Palapa: Jurnal Studi Keislaman dan Ilmu Pendidikan*, 5/1, 74-195.
- Munir, A. (2010). *Pendidikan karakter: membangun karakter anak sejak dari rumah*. Yogyakarta: Pustaka Insan Madani.
- Neuman, W. Lawrence. (2014). *Social Research Methods: Qualitative and Quantitative Approach* (7th ed). Edinburgh Gate Harlow: Pearson New International Edition.
- Novodvorsky, I. (1993). *Development of An Instrument to Assess Attitudes Toward Science*. Retrieved 5 October, 2018 from https://repository.arizona.edu/bitstream/handle/10150/186455/azu_td_9410657_sip1_m.pdf?sequence=1.
- Nunally, J. (1978). *Psychometric theory* (2nd ed). New York: McGraw Hill.
- Ogunleye, B. O. & Babajide, V. F. T. (2011). Generative Instructional Strategy Enhances Senior Secondary School Students' Achievement In Physics. *European Journal of Educational Studies* 3(3): 453-463.
- Purwanto. (2010). *Instrumen Penelitian Sosial dan Pendidikan: Pengembangan dan Pemanfaatan*. Yogyakarta: Pustaka Pelajar.
- Random Acts of Kindness. (2013). *Honesty and integrity*. The Random Acts of Kindness Foundation.
- Retnawati, H. 2017. *Validitas Reliabilitas dan Karakteristik Butir (Panduan untuk Peneliti, mahasiswa, dan Psikometrian)*. Yogyakarta: Parama Publishing.
- Robinson, J. P., Shaver, P. E., & Wrightsman, L. S. (1991). *Criteria for Scale Selection and Evaluation: In Measures of Personality and Social Psychological Attitudes*. San Diego, CA: Academic Press.

- Salahudin, A. (2013). *Pendidikan Karakter Berbasis Agama & Budaya Bangsa*. Bandung: Pustaka Setia.
- Samani, M. & Hariyanto. (2016). *Pendidikan karakter*. Bandung: PT Remaja Rosdakarya.
- Santyasa, I. W. (2014). *Asesmen dan Evaluasi Pembelajaran Fisika [Assessment and Evaluation of Physics Learning]*. Yogyakarta: Graha Ilmu.
- Sarah, S. (2018). *Penilaian Living Values dalam Pembelajaran Fisika*. Yogyakarta: K-Media.
- Sarah, S.; Prasetyo, Z. K.; Wilujeng, I. (2018). Development of Living Values (Responsibilities) Assessment Instruments in Learning Physics. *Advances in Social Science, Education and Humanities Research (ASSEHR)*, volume 267 5th Asia-Pacific Education Conference (AECON 2018).
- Sarah, S.; Prasetyo, Z. K.; Wilujeng, I. (2019). Development of Living Values (Honesty) Assessment Instrument in Learning Process. *International Journal of Instruction*, 12 (4): 443-458.
- Scheider, W. (1980). 27 Commandments for Physics Teachers. *The Physics Teacher*, 18 (1).
- Shalvi, S., Eldar, O., & Meyer, Y. B. (2012). Honesty Requires Time (and Lack of Justifications). *Psychological Science*, 23/10, 1264 –1270.
- Shihab, M. Q. (2007). *Secercah cahaya ilahi: hidup bersama al qur'an*. Bandung: Mizan.
- Sudaryono, Margono, G., Rahayu, W. (2013). *Pengembangan Instrumen Penelitian Pendidikan*. Yogyakarta: Graha Ilmu.
- Sugiyono. (1999). *Metode Penelitian Bisnis*, Bandung: Alfabeta.
- Suryabrata, Sumadi. 2005. *Pengembangan Alat Ukur Psikologis*, Yogyakarta: Andi.
- Tillman, D. (2004). *Living Values Activities for Young Adults*. Jakarta: Grasindo.

- _____. (2012b). *Living green values activities for children and young adults a special rio+20 edition*. Switzerland: Association for Living Values Education International (ALIVE).
- Usman, H. & Akbar, P. S. (2003). *Metodologi Penelitian Sosial*, Jakarta: Bumi Aksara.
- Woolfolk, A. E. & McCane, L. N. (1984). *Educational Psychology for teachers*. Englewood Cliffs, NJ.: Prentice Hall, In.
- Zuchdi, D., Kuntoro, S. A., & Prasetyo, Z. K. (2012). *Pendidikan Karakter: Konsep Dasar dan Implementasi di Perguruan Tinggi* [Character Building: Basic Concepts and Implementation in Higher Education] Yogyakarta: UNY Press.
- Zuchdi, D., Prasetya, Z. K., & Masruri, M. S. (2013). *Model Pendidikan Karakter Terintegrasi dalam Pembelajaran dan Pengembangan Kultur Sekolah* [Integrated Character Education Model in School Culture Learning and Development]. Yogyakarta: Multi Presindo.

BIODATA PENULIS



Rohmad, Lahir di Karanganyar 22 Desember 1966. Menyelesaikan Sekolah Dasar Negeri Kalioso III Gondangrejo tahun 1977, lulus Sekolah Menengah Pertama Negeri 1 Gemolong Sragen tahun 1981, lulus Pendidikan Guru Agama Negeri Surakarta tahun 1984, menyelesaikan Sarjana Lengkap Fakultas Tarbiyah Jurusan Pendidikan Agama Islam IAIN Sunan Kalijaga Yogyakarta tahun 1989 (Drs.), menyelesaikan Magister (S2) Pendidikan dan Evaluasi Pendidikan Program Pascasarjana Universitas Negeri Yogyakarta tahun 2001 (M. Pd.), menyelesaikan jenjang Doktor (S3) Studi Islam di Universitas Islam Negeri Sunan Kalijaga Yogyakarta tahun 2014.

Pengalaman mengajar: menjadi Dosen Luar Biasa di Fakultas Tarbiyah IAIN Sunan Kalijaga Yogyakarta tahun 1990 – 1992. Menjadi Dosen Tetap di Universitas Islam Negeri Profesor K.H. Saifuddin Zuhri Purwokerto sejak 1991 (Sebelumnya Fakultas Tarbiyah IAIN Sunan Kalijaga Yogyakarta di Purwokerto tahun 1964-1995, Fakultas Tarbiyah IAIN Walisongo Semarang di Purwokerto

tahun 1995-1997, Sekolah Tinggi Agama Islam Negeri Purwokerto tahun 1997-2015, Institut Agama Islam Negeri Purwokerto tahun 2015-2021, dan mulai Mei 2021 menjadi UIN Profesor K.H. Saifuddin Zuhri Purwokerto.

Beberapa Karyanya dalam bentuk buku antara lain: *Pengantar Statistika*, Yogyakarta: Kalimedia, 2015; *Pengembangan Instrumen Evaluasi dan Penelitian*, Yogyakarta: Kalimedia, 2017; *Evaluasi Program Full Day School di Madrasah Ibtidaiyah*, Purwokerto: Unsoed Press, 2019; *Pengembangan Instrumen Evaluasi Ranah Afektif Mata Pelajaran Akidah Akhlak di Madrasah Tsanawiyah*, Banyumas: Rizkuna, 2019. *Penilaian Hots: Pembelajaran PAI dan BP di SD*, Rizkuna, 2021.

Beberapa tulisan yang dipublikasikan di Jurnal Ilmiah antara lain:

- *The Evaluation of Curriculum Implementation on Islamic Higher Education in Indonesia.*
<http://www.ejournal.iainpurwokerto.ac.id/index.php/ijtimaiyya/article/view/929>
- *The Implementation of Authentic Assessment on 2013 Curriculum in Islamic Religious and Character Educational Learning at Purwokerto's High School.*
<http://www.ejournal.iainpurwokerto.ac.id/index.php/icms/article/view/2406>
- *Authentic Assessment of Affective Domain in Islamic Education and Character Building at A Curriculum 2013 Pilot Project Elementary School in Banyumas Regency Central Java Indonesia.*
<http://www.ejournal.iainpurwokerto.ac.id/index.php/icms/article/view/3028>
- *Evaluation of Islamic Full Day School Programs: Evaluation of Context, Input, Process, and Product of School Programs.*
<http://www.psychologyandeducation.net/pae/index.php/pae/article/view/2798>



Siti Sarah lahir di Purbalingga, pada tanggal 25 Mei 1982. Pendidikan S1 ditempuh pada prodi Pendidikan Fisika FMIPA Universitas Negeri Yogyakarta (2000-2005), jenjang S2 pada Magister Pendidikan Sains Konsentrasi Pendidikan Fisika Universitas Negeri Yogyakarta (2012-2014) dan jenjang S3 ditempuh pada Program Doktor Ilmu Pendidikan Konsentrasi Pendidikan Fisika Universitas Negeri Yogyakarta (2015-2019).

Pengalaman mengajar: dosen di Universitas Sains Al Qur'an, Wonosobo, Jawa Tengah (2009-2020). Mulai tahun 2020 hingga saat ini bekerja di Universitas Islam Negeri Profesor Kia Haji Saifuddin Zuhri, Purwokerto, Jawa Tengah.

Buku yang pernah diterbitkan di antaranya: *Physics Grade XA*, Yogyakarta: LP2IP (2009); *Fisika untuk SMK Kelas X*, Yogyakarta: LP2IP (2015); *Penilaian Living Values Dalam Pembelajaran Fisika*, Yogyakarta: K-Media (2018), *Model Pembelajaran Fisika Berbasis Potensi Lokal dan pedoman Penggunaannya: Upaya Mencapai Keterampilan Berpikir Kreatif Dan Living Values*, Yogyakarta: K-

Media (2019), dan *Fisika untuk SMK Kelas XI*, Yogyakarta: LP2IP (2020).

Beberapa karya berupa tulisan yang dipublikasikan antara lain: Keefektifan Pembelajaran Berbasis Potensi Lokal Dalam Pembelajaran Fisika SMA dalam Meningkatkan *Living Values* Siswa (2014), Pengembangan Perangkat Pembelajaran Berbasis Potensi Lokal Untuk Meningkatkan *Living Values* Peserta Didik SMA Di Kabupaten Wonosobo (2014), Penggunaan Modul Berbasis Inkuiri Untuk Meningkatkan Hasil Belajar dan Karakter Mandiri Siswa (2016), *Development Of Living Values (Responsibilities) Assessment Instryments In Learning Physics* (2018), *Integration Of Living Values Into Physics Learning Based On Local Potentials* (2018), Peningkatan Keterampilan Berpikir Kreatif Melalui Pembelajaran Fisika Berbasis Potensi Lokal (2018), Pandangan Filsafat Pragmatis John Dewey Dan Implikasinya Dalam Pendidikan Fisika (2018), Peningkatan Kerja Sinergis Melalui Pembelajaran Fisika Berbasis Potensi Lokal (2019), pemetaan Potensi Lokal Kabupaten Wonosobo Untuk Pembelajaran Fisika Sekolah Menengah Atas (SMA) (2019), *Development Of Living Values (Honesty) Assessment Instrument In Learning Process* (2019), *Problem-Solving Approach In Multiple Representations Of Qualitative And Quantitative Problems In Kinematics Motion* (2019), Pengembangan Instrumen Penilaian Living Values (Kerja Sinergi) Dalam Pembelajaran Fisika SMA (2019), *The Effectiveness Of Poe (Predict-Observe-Explain) Learning Model For Improving Student Analytical Skills* (2021), *Criticizing Higher Education Policy In Indonesia: Spiritual Elimination And Dehumanisation* (2021), *Android-Based Learning Media Development To Improve Student Learning Achievement* (2021). Semua tulisan yang dipublikasikan dapat dilihat pada link <https://scholar.google.co.id/citations?user=pvCiVKAAAAAJ&hl=id>