

Instructional Quality Questionnaire: Development of Instructional Quality Measurement Tools for Secondary Education Teachers

Instructional Quality Questionnaire: Pengembangan Alat Ukur Kualitas Instruksional Guru Jenjang Pendidikan Menengah

Megah Syifa Rinjani Sujarko¹
¹Fakultas Psikologi,
Universitas Padjadjaran, Indonesia
Email: megah15001@mail.unpad.ac.id

Airin Triwahyuni²
²Fakultas Psikologi,
Universitas Padjadjaran, Indonesia
Email: airin.triwahyuni@unpad.ac.id

Correspondence:

Megah Syifa Rinjani Sujarko
Fakultas Psikologi, Universitas Padjadjaran, Indonesia
Email: megah15001@mail.unpad.ac.id

Abstract

Instructional quality is a form of classroom teaching which is important to support the development of students' abilities in class. Instructional quality has three dimensions, that is a collection of teachers' teaching behaviors, namely classroom management, supportive climate, and cognitive activation. Measurements of instructional quality has begun to develop; however, no measurement scale has been constructed and used in Indonesia. This study discussed the adaptation of Instructional Quality Questionnaire with the aim of obtaining a measurement for instructional quality that is standardized and can be used in Indonesian population. The adaptation process carried out includes pre-condition stage, development stage and confirmation stage. A tryout of the instrument is carried out on 242 high school students. The final instructional quality scale consists of 18 items which measure all three dimensions with reliability ($r=0,886$) which indicates that the instrument is reliable. Collection of internal structure validity using confirmatory factor analysis shows that the validity construct of the instructional quality model is a good fit. This adaptation of instrument is expected to help the development of research related to instructional quality as an important construct in determining the increase in students' abilities and achievements, especially in Indonesia. Meanwhile, the construction of the items that is flexible to be used in any population and context is expected to facilitate the needs of various research.

Keyword : Instrument Adaptation, Instructional Quality, Teacher

Abstrak

Instructional quality atau kualitas instruksional guru merupakan bentuk dari classroom teaching yang penting untuk mendukung perkembangan kemampuan siswa di kelas. Instructional quality memiliki tiga dimensi berupa kumpulan perilaku guru ketika mengajar yaitu classroom management, supportive climate dan cognitive activation. Pengukuran kualitas instruksional telah mulai dikembangkan namun belum terdapat skala pengukuran yang telah dikonstruksikan dan digunakan di Indonesia. Penelitian ini membahas mengenai adaptasi Instructional Quality Questionnaire dengan tujuan untuk mendapatkan alat ukur instructional quality yang dapat digunakan pada populasi di Indonesia yang terstandarisasi. Proses adaptasi yang dilakukan meliputi tahap pre-kondisi, pengembangan dan konfirmasi. Uji coba alat ukur dilakukan kepada 242 siswa SMA/SMK/ sederajat. Alat ukur final instructional quality terdiri dari 18 item yang mengukur ketiga dimensinya dengan reliabilitas ($r=0,886$) yang mengindikasikan alat ukur reliabel. Pengumpulan bukti validitas struktur internal dengan menggunakan confirmatory factor analysis menunjukkan bahwa model instructional quality memiliki hasil validitas konstruk baik (good fit). Pengembangan terhadap alat ukur instructional quality ini diharapkan dapat membantu perkembangan penelitian terkait konstruk perilaku instruksional sebagai konstruk yang penting dalam menentukan peningkatan kemampuan dan pencapaian siswa khususnya di Indonesia. Penyusunan item alat ukur yang fleksibel untuk dapat digunakan pada berbagai populasi dan konteks juga diharapkan dapat memfasilitasi kebutuhan beragam tujuan penelitian.

Kata Kunci : Adaptasi Alat Ukur, Guru, Kualitas Instruksional

Copyright (c) 2024 Megah Syifa Rinjani Sujarko & Airin Triwahyuni

Received 2023-10-09

Revised 2023-11-19

Accepted 2024-02-19



LATAR BELAKANG

Proses belajar mengajar di kelas (*classroom teaching*) merupakan metode pengajaran yang paling umum digunakan di sekolah/institusi pendidikan lain (Gage & Berliner, 1998). Komponen penting dari *classroom teaching* yaitu adanya guru yang mampu melakukan praktik instruksional secara efektif yang disebut pula sebagai *instructional quality* (Bellens et al., 2019; Fauth et al., 2014; Kunter & Baumert, 2006; Praetorius et al., 2018). *Instructional quality* merupakan kumpulan berbagai bentuk perilaku yang digunakan guru untuk menciptakan situasi belajar yang menantang dan adaptif, serta membimbing siswa secara hati-hati selama proses belajar (Kunter & Baumert, 2006).

Pengembangan praktik instruksional yang efektif merupakan dasar dari segala upaya untuk meningkatkan pencapaian siswa (Brophy, 1986; Yilmaz et al., 2017), motivasi belajar (Sigala, 2016; Yu & Singh, 2018), minat (Förtsch et al., 2017) serta keyakinan diri siswa dalam belajar (Sakiz et al., 2012). Artinya untuk dapat mengembangkan kemampuan siswa di sekolah, guru terlebih dahulu perlu memiliki kualitas instruksional atau kemampuan mengajar yang baik sehingga muncul kebutuhan untuk dapat mengetahui kualitas instruksional seperti apa yang perlu dimiliki oleh seorang guru. Untuk dapat mengetahui hal tersebut, pertama kita perlu mampu mengukur kualitas dari praktik instruksional. Dalam penelitian ini, akan dilakukan pengembangan alat ukur untuk mengukur kualitas ajar guru yaitu alat ukur *Instructional Quality Questionnaire*.

Pengembangan *Instructional Quality Questionnaire* dilakukan dengan mengadaptasi alat ukur *Instructional Quality Questionnaire (for Flanders and Germany students)* yang dikonstruksi oleh Bellens et al., (2019). Adaptasi alat ukur dilakukan karena diperlukan alat ukur yang mampu mengukur tingkat *instructional quality* dengan memberikan penilaian terhadap perilaku guru dalam proses belajar mengajar. Di Indonesia khususnya, belum terdapat alat ukur yang diperuntukkan untuk mengukur konstruk *instructional quality*. Sebelumnya, cukup banyak penelitian-penelitian yang dilakukan untuk meneliti mengenai *instructional quality* serta hubungannya dengan berbagai konstruk atau aspek lain dalam pendidikan (Baumert et al., 2013; Guill et al., 2020; Kunter et al., 2011; Yilmaz et al., 2017; Yu & Singh, 2016).

Namun masih sedikit penelitian yang dilakukan di negara/populasi non-Barat (Nilsen & Gustafsson, 2016). Selain itu, penelitian-penelitian sebelumnya menggunakan *instructional quality questionnaire* untuk meneliti populasi siswa tingkat sekolah dasar (Bellens et al., 2019; Fauth et al., 2014) dan sekolah menengah pertama (Kunter & Baumert, 2006) pada konteks mata pelajaran tertentu. Instruksi serta *item* pada alat ukur dibuat secara spesifik untuk digunakan pada populasi dan konteks yang disasar. Pada penelitian ini, instruksi serta *item* alat ukur akan dibuat menjadi lebih fleksibel sehingga alat ukur dapat disesuaikan dan digunakan untuk populasi dan konteks (belajar mengajar) yang secara umum dapat mencakup berbagai tujuan penelitian. Dengan disusunnya alat ukur ini diharapkan dapat membantu perkembangan penelitian-penelitian lain terkait kualitas instruksional guru di Indonesia.

Instructional Quality (Kualitas Instruksional)

Suatu praktik pendidikan dilakukan dengan memiliki tujuan yaitu untuk memfasilitasi terjadinya proses belajar dan perkembangan dan pencapaian diri siswa baik secara kognitif maupun afektif. Perkembangan siswa lewat proses pendidikan dapat terjadi ketika pendidikan yang diberikan merupakan pendidikan yang efektif.

Creemers dan Kyriakides (2008, dalam Nilsen & Gustafsson, 2016) menyusun model dinamis efektivitas pendidikan. Model tersebut merupakan kerangka konseptual yang menggambarkan penentu-penentu (determinan) dari hasil belajar siswa (*student outcomes*). Terdapat 4 level determinan pada kerangka tersebut yaitu 1) level nasional, 2) level sekolah, 3) level kelas dan 4) level siswa. *Classroom teaching* yang merupakan metode pengajaran paling umum dilakukan (Gage & Berliner, 1998) terjadi pada level ketiga yaitu level kelas. Level kelas memiliki dua aspek yang penting bagi hasil dari pendidikan berupa *teacher quality* dan *instructional quality*. Kualitas guru dan kualitas instruksional merupakan penentu utama dari hasil belajar siswa (Nilsen & Gustafsson, 2016). Pada level inilah proses belajar mengajar terjadi.

Kualitas instruksional guru (*instructional quality*) merupakan konstruk yang mencerminkan aspek-aspek guru dalam mengajar yang diketahui berhubungan positif dengan pencapaian siswa, baik secara kognitif maupun afektif (Bellens et al., 2019). Dilihat dari berbagai praktik instruksional guru, ditemukan kesepakatan mengenai 3 dimensi perilaku pembentuk konstruk *instructional quality*. Ketiga dimensi tersebut yaitu 1) *classroom management*, 2) *supportive climate* dan 3) *cognitive activation* (Bellens et al., 2019). Ketiga dimensi ini merupakan kumpulan perilaku-perilaku yang ditunjukkan oleh guru selama proses belajar mengajar. Perilaku-perilaku yang dinilai pada *instructional quality* ini merupakan perilaku yang penting untuk membangun pengetahuan aktif dan mandiri pada siswa di kelas (Kunter & Baumert, 2006).

Classroom management merupakan perilaku-perilaku ajar yang berfokus pada penegakkan aturan dan prosedur kelas, mengatasi gangguan yang terjadi, serta kemampuan transisi yang mulus. Dimensi ini merupakan keadaan awal atau prasyarat yang perlu dilaksanakan agar proses belajar mengajar dapat terjadi. *Supportive climate* merupakan hubungan antar guru dan siswa yang suportif, perilaku pemberian umpan balik dari guru yang positif dan konstruktif, pendekatan positif terhadap kesalahan maupun kesalahpahaman siswa, pemberian dukungan individual serta perilaku peduli pada siswa. *Cognitive activation* merupakan praktik pedagogis yang digunakan oleh guru untuk mendorong keterlibatan siswa dalam proses berpikir tingkat tinggi serta usaha meningkatkan keterlibatan siswa secara kognitif yang dapat menghasilkan pembelajaran. Di dalamnya termasuk juga pemilihan bentuk masalah/tugas yang diberikan kepada siswa serta metode penyampaiannya.

Prestasi atau pencapaian siswa hasil dari proses belajar seringkali digunakan sebagai tolak ukur kesuksesan maupun kegagalan dari suatu praktik pendidikan. Model dinamis efektivitas pendidikan dari Creemers dan Kyriakides

(2008, dalam Nilsen & Gustafsson, 2016) menunjukkan bahwa *instructional quality* dapat memengaruhi hasil belajar siswa, baik secara kognitif (prestasi/*achievement*) maupun afektif (motivasi, emosi, dan lain sebagainya). Sehingga tingkat *instructional quality* yang baik dari guru diharapkan dapat menghasilkan hasil belajar yang baik pula bagi siswa.

Perkembangan Pengukuran Instructional Quality

Pengukuran *instructional quality* telah mulai dikembangkan oleh Kunter & Baumert (2006) yang sebelumnya telah mengkaji konsep *instructional quality* kemudian mengembangkan *instructional quality questionnaire teacher's rating* dan *student's rating*. Pengembangan alat ukur tersebut dimaksudkan untuk menentukan cara pengukuran kualitas instruksional yang paling tepat melalui guru atau siswa. Alat ukur tersebut dikonstruksikan *class specific* yaitu untuk kelas matematika. Metode lain yang dikembangkan sebagai pengukuran *instructional quality* yaitu menggunakan metode observasi disamping metode lainnya (Lipowsky et al., 2009). Berbagai penelitian dan pengembangan mengenai konsep *instructional quality* dan pengukurannya dilakukan, berikut pula kajian mengenai definisi dari dimensi-dimensi *instructional quality*. Praetorius et al. (2014) melakukan kajian terhadap berbagai definisi tersebut dan menemukan definisi *classroom management* secara umum hampir sama pada berbagai penelitian lintas budaya, namun operasionalisasi dimensi *supportive climate* dan *cognitive activation* masih memiliki perbedaan akibat dimensinya yang *content* dan *setting specific* (tergantung kepada situasi spesifik yang terjadi pada populasi tertentu).

Berdasarkan kajian literatur, peneliti memilih untuk mengadaptasi alat ukur yang digunakan oleh Bellens et al. (2019). Pada penelitiannya, Bellens et al. (2019) menggunakan dua skala pengukuran berupa kuesioner *student's rating* yang ditujukan pada tiga populasi berbeda yaitu siswa SMP di Belgia (Flanders), Jerman dan Norwegia. Penelitiannya menghasilkan dua skala pengukuran yaitu *instructional quality questionnaire for Flanders and Germany students* dan *for Norway students*. Peneliti memilih untuk mengadaptasi skala pengukuran *instructional quality for Flanders and Germany students* dengan alasan (1) definisi dari konstruk dan dimensi yang digunakan pada penelitian umum (tidak *content/setting specific* dan dapat diterapkan pada aktivitas belajar di Indonesia, (2) skala pengukuran dimensi *supportive climate* yang digunakan pada skala Flanders dan Jerman mengukur *support* guru dari hubungan antara guru-siswa sedangkan pada skala Norwegia hal tersebut diukur melalui perilaku *support* yang guru berikan terkait konten pembelajaran di kelas.

Peneliti menilai bahwa dibandingkan dengan skala Norwegia, skala Flanders dan Jerman lebih cocok dengan budaya dan kegiatan belajar yang terjadi di Indonesia karena melihat interaksi antara guru dan siswa di sekolah terjadi tidak hanya di dalam kelas dan terkait konten pembelajaran saja namun lebih luas meliputi juga kegiatan-kegiatan luar kelas. Kuesioner *instructional quality for Flanders and Germany students* diuji coba menggunakan data dari *Trends in International Mathematics and Science Study 2015* berupa data siswa kelas 4 dari Flanders dan Jerman. Hasil uji coba menunjukkan bahwa alat ukur valid dan reliabel dengan modifikasi. Hasil uji validitas menggunakan CFA pada alat ukur dapat dilihat masing-masing pada Tabel 1.

Tabel 1. Hasil CFA *Instructional Quality for Flanders and Germany Students* (Tanpa Modifikasi)
Sumber: Bellens et al. (2019); Hu & Bentler (1999)*; Hair et al.**

Skala Alat Ukur	Parameter fit	Kriteria Fit*	Output	Keterangan
Flanders Scale	X^2/df	$\leq 3,00^{**}$	10,73	Tidak fit
	RMSEA	$\leq 0,05$	0,04	Fit
	CFI	$\geq 0,90$	0,83	Tidak fit
	SRMR <i>within class</i>	$\leq 0,09$	0,08	Fit
Germany Scale	X^2/df	$\leq 3,00^{**}$	9,65	Tidak fit
	RMSEA	$\leq 0,05$	0,05	Fit
	CFI	$\geq 0,90$	0,85	Tidak fit
	SRMR <i>within class</i>	$\leq 0,09$	0,08	Fit

METODE PENELITIAN

Alat Ukur Penelitian

Adaptasi alat ukur dilakukan terhadap *Instructional Quality Questionnaire for Flanders and Germany Students* dari Bellens et al. (2019). Alat ukur asli terdiri dari 17 item dengan rincian item per-dimensi serta indikatornya dapat dilihat pada Tabel 2. Skala pengukuran yang tersedia pada alat ukur asli yaitu berupa skala *likert* yang terdiri dari 4 pilihan jawaban: *never or hardly ever* (0), *some lessons* (1), *most lessons* (2) dan *every lesson* (3) untuk item-item yang mengukur dimensi *classroom management*, dan *disagree a lot* (0), *disagree a little* (1), *agree a little* (2) dan *agree a lot* (3) untuk item-item yang

mengukur dimensi *supportive climate* dan *cognitive activation*. Pengisian alat ukur dilakukan dengan partisipan diminta untuk memilih satu angka pada skala yang menunjukkan keadaan dirinya pada saat mengikuti pembelajaran pada mata pelajaran tertentu. Skala pengukuran yang digunakan pada alat ukur *instructional quality* setelah hasil adaptasi yaitu berupa 4 pilihan respon sangat tidak setuju (skor 1), tidak setuju (2), setuju (3) dan sangat setuju (4). Perubahan skala dari dua bentuk pilihan menjadi satu dilakukan berdasarkan hasil rekomendasi pada proses review konstruk dengan mengubah beberapa struktur kalimat pada item (lihat Tabel 3). Perubahan struktur kalimat yang dilakukan yaitu menambahkan subjek pada pernyataan

yang diberikan pada item (“terdapat kebisingan dan gangguan”) menjadi “suasana kelas berisik saat guru mengajar”). Skala tersebut kemudian digunakan saat penelaahan oleh ahli dan *pilot test*. Beberapa item kembali direvisi berdasarkan hasil *expert review* dengan mengubah item agar lebih menekankan pada perilaku yang dilakukan guru di kelas dibandingkan keadaan kelas itu sendiri (“suasana kelas berisik saat guru mengajar” menjadi “guru mampu menenangkan suasana kelas yang berisik”). Pada

instruksi pengisian alat ukur, diberikan keterangan mengenai mata pelajaran spesifik yang akan diukur beserta guru yang mengajar pada mata pelajaran tersebut sehingga partisipan dapat membayangkan dan mengingat situasi belajarnya pada pelajaran tersebut. Hasil *expert review* dan *pilot test* menunjukkan bahwa dengan digunakannya skala tersebut, item alat ukur tetap mampu mengukur dimensi yang diharapkan.

Tabel 2. Kisi-kisi Alat Ukur *Instructional Quality* (for Flanders and Germany Students)
Sumber: Bellens et al. (2019)*

*Dimensi	*Jumlah Item	Indikator	Nomor Item	*Contoh Item dari Alat Ukur Asli	Contoh Item Hasil Adaptasi
Classroom management (CM)	5	Guru mampu mengatasi gangguan yang terjadi di kelas	2	<i>There is noise and disorder</i>	Guru mampu menenangkan suasana kelas yang berisik
		Guru mampu membuat siswa mengikuti aturan dan prosedur di kelas	1, 4	<i>Students don't listen to what the teacher say</i>	Guru mampu membuat siswa mendengarkan materi pelajaran yang sedang disampaikan
		Guru mampu mengoptimalkan waktu yang tersedia	3, 5	<i>Students don't start working for a long time after the lesson begins</i>	Guru mampu dengan cepat membuat siswa tenang sebelum memulai pelajaran
Supportive climate (SC)	5	Guru menunjukkan kepedulian kepada siswa	7, 9	<i>Our mathematics teacher cares about me</i>	Guru menunjukkan sikap peduli kepada saya
		Guru memberikan dukungan kepada siswa saat menghadapi permasalahan	6, 8, 10	<i>Our mathematics teacher is nice to me even when I make a mistake</i>	Guru memberi dukungan kepada saya bahkan ketika saya membuat kesalahan saat mengerjakan latihan/tugas
Cognitive activation (CA)	7	Guru memberikan tugas yang mendorong siswa untuk berpikir	11, 13, 14, 16	<i>Our mathematics teacher gives us tasks I like to think about</i>	Guru memberikan pertanyaan mengenai materi pelajaran yang harus saya pikirkan dengan sungguh-sungguh
		Guru memastikan siswa dapat memahami tugas/materi yang diberikan	12, 15, 17	<i>Our mathematics teacher asks me what I have understood and what I haven't</i>	Guru menanyakan materi apa yang sudah saya pahami dan apa yang belum

Prosedur Penelitian

Prosedur Adaptasi Alat Ukur

Proses adaptasi alat ukur dilakukan dengan mengikuti tahapan adaptasi dari *International Test Commission* (2017) yang secara umum terdiri dari tahap pre-kondisi, tahap pengembangan dan tahap konfirmasi alat ukur. Tahap pre-kondisi dilakukan dengan tujuan mendapatkan izin untuk mengadaptasi dan mereview konstruk alat ukur pada budaya tempat tujuan adaptasi. Tahap pengembangan bertujuan untuk memastikan proses adaptasi dapat menghasilkan alat ukur yang dapat dipahami serta sesuai dengan bahasa dan budaya populasi target. Sementara tahap konfirmasi dilakukan untuk mendapatkan bukti bahwa hasil adaptasi alat ukur sudah dapat digunakan pada populasi target dan dapat mengukur konstruk atau dimensi yang diharapkan.

Pada proses adaptasi alat ukur ini, tahap pre-kondisi yang dilakukan yaitu menghubungi peneliti sebelumnya untuk mendapatkan izin melakukan adaptasi terhadap alat ukur *instructional quality*. Setelah memperoleh izin adaptasi, dilakukan tahap selanjutnya yaitu tahap pengembangan alat ukur yang secara garis besar terdiri dari proses *translation*, *review* dan *evaluation*.

1. Proses *forward translation*. Proses translasi dilakukan dengan teknik *forward-backward translation*. Proses translasi pertama yaitu *forward translation* dimana alat ukur diterjemahkan dari bahasa asli (Inggris) ke Bahasa Indonesia. Proses translasi *forward* dilakukan oleh 2 orang translator dengan kualifikasi translator 1: (1) aktif bekerja sebagai translator, (2) merupakan seorang dosen program studi Bahasa Inggris, (3) memiliki sertifikasi

- sebagai translator profesional, (4) anggota muda Himpunan Penerjemah Indonesia; dan kualifikasi translator 2: (1) aktif bekerja sebagai translator, (2) memiliki latar belakang pendidikan di bidang psikologi, dan (3) anggota muda Himpunan Penerjemah Indonesia.
2. Proses *review* konstruk. *Review* dilakukan oleh 3 orang rekan sejawat peneliti terhadap hasil terjemahan. Telaah yang dilakukan antara lain mengenai (1) ketepatan terjemahan dari bahasa asli alat ukur ke Bahasa Indonesia, (2) kesesuaian budaya pada *item* alat ukur dengan keadaan populasi di Indonesia, (3) telaah konten serta (4) penyesuaian bahasa (redaksi). Kualifikasi dari individu yang melakukan *review* yaitu: (1) merupakan Sarjana Psikologi, (2) mahasiswa aktif magister profesi Psikologi, (3) memiliki pengalaman dalam penyusunan alat ukur psikologi. Proses *review* dilakukan dengan cara *reviewer* menelaah alat ukur asli dan hasil translasi *forward* dan memberikan koreksi dan rekomendasi dengan mempertimbangkan latar belakang teori dari konstruk/dimensi yang diukur. Selain rekomendasi terkait *item* alat ukur, diperoleh juga rekomendasi terkait perubahan pada skala yang digunakan pada alat ukur adaptasi.
 3. Proses *backward translation*. Hasil *review* kemudian diterjemahkan kembali ke bahasa asli yaitu bahasa Inggris lewat proses translasi *backward*. Translasi *backward* dilakukan oleh dua orang translator dengan kualifikasi translator 1: (1) aktif bekerja sebagai translator, (2) memiliki latar belakang pendidikan di bidang psikologi; dan kualifikasi translator 2: (1) aktif bekerja sebagai translator, (2) anggota muda Himpunan Penerjemah Indonesia.
 4. Proses *expert review*. Hasil dari proses translasi dan *peer review* diberikan kepada seorang ahli untuk dilakukan *expert review*. *Expert* yang melakukan *review* memiliki kualifikasi antara lain (1) dosen tetap di Sekolah Pascasarjana Universitas Pakuan, Bogor, (2) memiliki latar belakang pendidikan S1 psikologi dan S2 serta S3 di bidang manajemen pendidikan, dan (3) direktur dari sekolah swasta di Bogor. Poin-poin yang ditelaah pada proses *review* ini antara lain (1) kesesuaian definisi antara konstruk, dimensi dan *item*, (2) ketepatan redaksi *item* dan (3) relevansinya untuk mengukur dimensi dan konstruk yang dimaksud. Proses *review* dilakukan dengan cara pengisian *form expert judgement* rancangan alat ukur yang telah melalui proses adaptasi sebelumnya (translasi *forward-backward* dan *peer review*). *Expert* memberikan

- review* berupa koreksi, masukan maupun rekomendasi terhadap *item* alat ukur dengan memerhatikan latar belakang konstruk/dimensi yang diukur. Hasil dari *review* dianalisa secara kualitatif dengan mengelompokkan (*coding*) data yang diperoleh yang dapat dilihat pada Tabel 3. Data hasil *expert review* lengkap terlampir pada Appendix 4.
5. Proses evaluasi/sintesis. Hasil dari proses adaptasi kemudian dievaluasi oleh peneliti yang menghasilkan beberapa penyesuaian redaksi kalimat pada *item* 1, 3, 4, 5, 6, 9, 10, 11, 12, 16, dan 17. Selain itu pada *item* nomor 2 dilakukan penyesuaian dengan membagi *item* menjadi dua sehingga total *item* alat ukur menjadi sebanyak 18 *item* dimana nomor *item* dicantumkan menjadi 2a dan 2b.

Selanjutnya dilakukan tahap konfirmasi dengan metode *pilot testing* untuk menguji hasil adaptasi alat ukur yang dilakukan dengan menggunakan metode *cognitive interview*. *Interview* dilaksanakan secara daring melalui *video call* terhadap 2 orang partisipan secara bersamaan. Alat ukur hasil dari proses adaptasi sejumlah 18 *item* digunakan pada *pilot test* kepada dua orang siswa SMA. *Cognitive interview* ini dilakukan sebagai pengumpulan bukti *content validity*. *Cognitive interview* dilakukan dengan metode wawancara *evidence based on response process* menggunakan *guideline interview* yang mengacu pada proses kognitif dari Conrad & Blair (dalam Willis, 2015) yaitu 1) *understanding the intent of the question* (apakah Anda memahami maksud dari pernyataan pada soal ini?); 2) *performing the primary task* (hal apa yang Anda pikirkan ketika membaca pernyataan pada soal ini?); dan 3) *formatting and mapping response* (dari soal-soal nomor sekian hingga sekian, menurut anda pernyataan-pernyataan ini menanyakan anda mengenai hal apa?). *Cognitive interview* dilakukan dengan tujuan menggali pemahaman partisipan terhadap persoalan, gambaran proses kognitif yang terjadi ketika memberikan jawaban, serta pemetaan jawaban partisipan terhadap alat ukur.

Hasil dari wawancara digunakan untuk melakukan penyesuaian *item* pada alat ukur sebelum dilakukan uji coba. Analisis hasil wawancara dilakukan secara kualitatif dengan melakukan pengelompokkan (*coding*) kata kunci dari jawaban partisipan yang sesuai dengan maksud yang dituju tiap *item*. Hasil *coding* jawaban partisipan *cognitive interview* dapat dilihat pada Tabel 3 dan data hasil wawancara dapat dilihat pada Appendix 5. Hasil dari *cognitive interview* menghasilkan penyesuaian kalimat pada *item* nomor 6, 7, 8, 10, 13, 15, 16, dan 17.

Tabel 3. Hasil Analisis Expert Review dan Cognitive Interview

Dimensi	Nomor Item yang Diperbaiki	Coding Perbaikan	Contoh Item Sebelum Perbaikan	Contoh Item Setelah Perbaikan
Expert Review				
Classroom management	1, 2a, 2b, 3, 4, 5	Konten	Siswa tidak menyimak apa yang disampaikan oleh gurunya	Guru mampu membuat siswa mendengarkan materi pelajaran yang sedang disampaikan

<i>Supportive climate</i>	6, 9, 10,	Pemahaman	Guru saya bersikap baik kepada saya bahkan ketika saya membuat kesalahan	Guru memberi dukungan kepada saya bahkan ketika saya membuat kesalahan
<i>Cognitive activation</i>	12, 16	Pemahaman	Guru bertanya kepada saya tentang apa yang sudah saya pahami dan apa yang belum	Guru menanyakan materi apa yang sudah saya pahami dan apa yang belum
	11, 17	Konten	Guru ingin saya mampu menjelaskan jawaban saya	Guru meminta saya menjelaskan jawaban saya
Cognitive Interview				
<i>Supportive climate</i>	6, 7, 8	Pemahaman	Guru memberi dukungan kepada saya bahkan ketika saya membuat kesalahan	Guru memberi dukungan kepada saya bahkan ketika saya membuat kesalahan saat mengerjakan latihan/tugas
	10	Proses kognitif	Guru saya yakin bahwa saya dapat menyelesaikan tugas yang sulit	Guru saya menunjukkan rasa percaya bahwa saya dapat menyelesaikan tugas yang sulit
<i>Cognitive activation</i>	13, 15, 16, 17	Pemahaman	Guru meminta saya menjelaskan jawaban saya	Guru meminta saya untuk dapat menjelaskan jawaban dari tugas yang saya kerjakan
	16	Proses kognitif	Guru memberi tugas yang membuat saya terdorong untuk menyelesaikannya meskipun sulit	Guru memberi tugas yang membuat saya termotivasi untuk menyelesaikannya walaupun sulit

Partisipan Penelitian

Setelah diadaptasi, alat ukur diuji cobakan kepada populasi siswa SMA/SMK/ sederajat. Uji coba alat tes dilakukan secara daring menggunakan *Google Form*. Uji coba alat tes dilakukan kepada 242 partisipan siswa siswi SMA/SMK/ sederajat yang terdiri dari 115 siswa SMA, 44 siswa SMK, 82 siswa MA dan 1 lainnya. Karakteristik partisipan beragam yaitu terdiri dari level kelas 10 (N=54), 11 (N=84) dan 12 (N=104) serta jenis kelamin laki-laki (N=46) dan perempuan (N=196). Teknik *sampling* yang digunakan yaitu *convenient sampling* untuk menjangkau subyek dengan kriteria yang diinginkan. Partisipasi dalam pengujian alat ukur bersifat sukarela dimana partisipan diberikan penjelasan mengenai tujuan dari uji coba alat ukur serta hak sebagai partisipan termasuk di dalamnya jaminan bahwa data yang diberikan dalam uji coba akan dijaga kerahasiaannya, penjelasan mengenai pemanfaatan data serta hak bagi partisipan untuk menolak/berhenti mengikuti kegiatan uji coba alat ukur kapanpun partisipan inginkan. Partisipan kemudian mengisi lembar *informed consent* yang menjelaskan kepada partisipan mengenai gambaran kegiatan uji coba alat ukur. Partisipan yang telah memahami informasi yang diberikan kemudian menunjukkan persetujuannya untuk berpartisipasi dalam penelitian dengan menandatangani pernyataannya pada lembar kesediaan.

Analisis Data

Analisis dilakukan untuk mendapatkan properti psikometri alat ukur *instructional quality* dan mengkonfirmasi model pengukuran dengan data populasi yang dituju yaitu siswa SMA sederajat. Data hasil uji coba alat ukur digunakan

untuk menganalisis apakah *item-item* alat ukur telah memenuhi persyaratan psikometris untuk diikutsertakan sebagai bagian dari alat ukur. Analisis data yang dilakukan pertama yaitu menguji daya beda *item* (*item discrimination*) yang dilakukan dengan pengujian *item total correlation* menggunakan program *IBM SPSS Statistic 25*. Daya diskriminasi *item* dilihat dari nilai koefisien *corrected total-item correlation*. Kemudian dilakukan pula analisis *item* berupa uji validitas konstruk menggunakan analisis *Confirmatory Factor Analysis* (CFA) dan uji reliabilitas *internal consistency* menggunakan *Cronbach's coefficient alpha*. Analisa CFA dengan estimasi *maximum likelihood* dilakukan menggunakan program perangkat lunak *Lisrel version 9.30* sedangkan analisa reliabilitas dilakukan menggunakan *IBM SPSS Statistic 25*.

HASIL PENELITIAN

Alat ukur *instructional quality* hasil proses adaptasi terdiri dari 18 *item* dengan komposisi *item* pada masing-masing dimensi yaitu *classroom management* 6 *item*, *supportive climate* 5 *item* dan *cognitive activation* 7 *item*. Dilakukan uji reliabilitas terhadap data hasil uji coba untuk mengetahui reliabilitas awal dari alat ukur. Pengujian reliabilitas *internal consistency* dilakukan menggunakan uji *Cronbach's alpha*. Nilai reliabilitas alat ukur *instructional quality* dan ketiga dimensinya disajikan pada Tabel 4. Dapat dilihat bahwa alat ukur *instructional quality* dan dimensi-dimensinya memiliki nilai reliabilitas $\geq 0,70$ (Kaplan & Saccuzzo, 2013) sehingga dapat dikatakan bahwa alat ukur sudah reliabel.

Tabel 4. Reliabilitas Alat Ukur Instructional Quality Setelah Adaptasi

Konstruk / Dimensi	Jumlah Item	Reliabilitas	Keterangan
Instructional quality	18	0,886	Reliabel
Dimensi Classroom management	6	0,857	Reliabel
Dimensi Supportive climate	5	0,829	Reliabel
Dimensi Cognitive activation	7	0,755	Reliabel

Analisis *item discrimination* dilakukan dengan melihat nilai koefisien *corrected total-item correlation*. Hasil pengujian *item discrimination* dapat dilihat pada Tabel 5. Kriteria pemilihan *item* berdasarkan koefisien *item-total correlation* dengan nilai $r_{ix} \geq 0,30$ memiliki daya diskriminasi yang tinggi, sedangkan *item* dengan $r_{ix} < 0,30$ memiliki daya diskriminasi rendah (Azwar, 2009; Ebel, 1967). Dari Tabel 6.1 dapat dilihat

bahwa *item* 14 memiliki nilai koefisien korelasi kurang dari 0.30 ($r_{ix} = 0.127$) sehingga terdapat indikasi bahwa *item* 14 kurang mampu untuk membedakan kemampuan guru untuk menstimulai siswa secara kognitif. Hal ini dijadikan pertimbangan untuk melakukan penyesuaian *item* untuk mendapatkan alat ukur final yang siap digunakan.

Tabel 5. Daya Diskriminasi Item Alat Ukur

Dimensi	Kode Item*	Corrected Item-Total Correlation
Classroom management	CM1	0.548
	CM2a	0.526
	CM2b	0.589
	CM3	0.533
	CM4	0.566
Supportive climate	CM5	0.566
	SC6	0.601
	SC7	0.578
	SC8	0.612
	SC9	0.531
Cognitive activation	SC10	0.611
	CA11	0.373
	CA12	0.520
	CA13	0.419
	CA14	0.127
	CA15	0.570
	CA16	0.585
CA17	0.496	

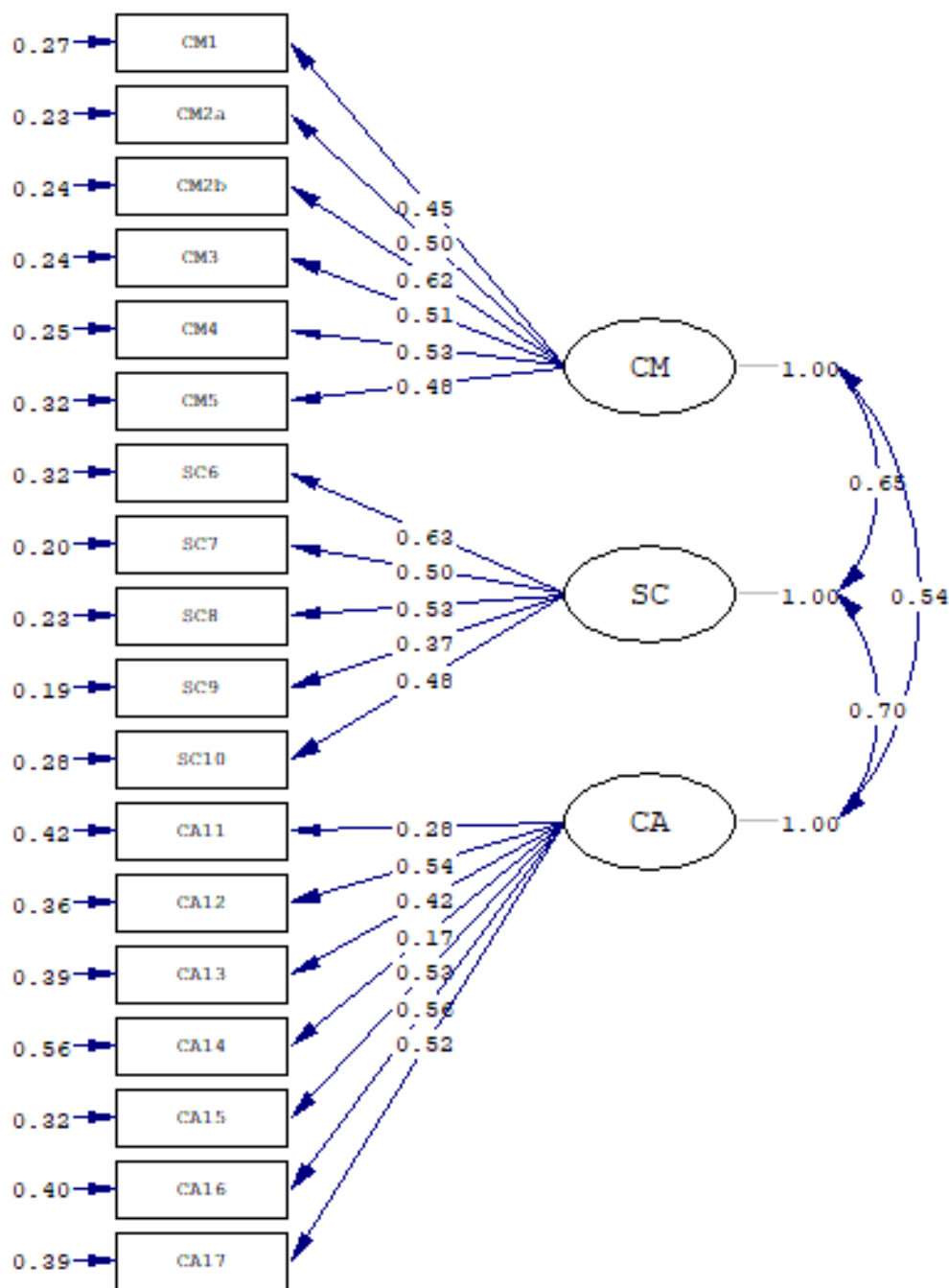
*CM: classroom management; SC: supportive climate; CA: cognitive activation

Pengumpulan bukti validitas struktur internal menggunakan metode CFA dilakukan terhadap alat ukur *instructional quality* hasil adaptasi. Diperoleh nilai *degree of freedom* yaitu 132. Model pengukuran dari alat ukur disajikan pada Gambar 1. Gambar 1 menunjukkan bahwa terdapat dua

buah *item* yang memiliki *loading factor* kurang dari 0,4 (Field, 2009) yaitu *item* CA11 (“guru memberi tugas yang membuat saya perlu berpikir dengan sungguh-sungguh”), dan *item* CA14 (“guru memberi tugas yang pada awalnya terlihat sulit”).

Tabel 6. Parameter Fit (Hair et al., 2019)* Model Instructional Quality

Parameter fit	Kriteria fit*	Output	Keterangan
<i>p-value</i>	>0,05	0,000	Tidak Fit
X^2/df	$\leq 3,00$	2,20	Fit
RMSEA	$\leq 0,08$	0,071	Fit
CFI	$\geq 0,95$	0,96	Fit
SRMR	$\leq 0,10$	0,063	Fit



Gambar 1. Model Pengukuran Alat Ukur Instructional Quality

PEMBAHASAN

Tabel 6 menyajikan hasil parameter fit dari model *instructional quality*. Dapat dilihat bahwa masih terdapat parameter fit yang belum mencapai kriteria dari *Goodness of fit* yang ditetapkan yaitu pada parameter fit *p-value*. Nilai *Chi-square* belum memenuhi kriteria fit ($X^2 = 290,43$; $p\text{-value} \leq 0,05$). Perhitungan *chi-square* yang melibatkan jumlah data (N) dapat menyebabkan sulitnya kriteria *fit* untuk dapat dicapai. Ketika jumlah data ditambah, nilai *chi-square* akan meningkat. Semakin besar *chi-square* maka *p-value* semakin kecil dan kriteria *fit* semakin sulit untuk dicapai (Hair et al., 2019). Maka rasio dari *chi-square* terhadap *degree of freedom*

(X^2/df) menjadi alternatif untuk dapat menilai kriteria *fit* (Hooper et al., 2008). Pada alat ukur hasil adaptasi diperoleh rasio *chi-square* dengan *df* yaitu $\leq 3,00$ (Hair et al., 2019) yang menunjukkan bahwa model memenuhi kriteria *fit*. Untuk melengkapi nilai *chi-square*, dilihat parameter fit lain yaitu RMSEA yang merupakan fungsi langsung dari *chi-square* dengan koreksi terhadap dampak jumlah sampel terhadap model (Umar & Nisa, 2020). RMSEA yang diperoleh model sebesar 0,071 dimana parameter telah memenuhi kriteria *goodness of fit* (lihat Tabel 6). Selain RMSEA, parameter lain juga nampak telah memenuhi kriteria *fit* (CFI = 0,960; SRMR = 0,063). Artinya, model dapat dikatakan valid.

Tabel 7. Hasil Analisis Item Alat Ukur Hasil Adaptasi

Dimensi	Jumlah Item	Reliabilitas Chronbach Alpha	Kode Item	Loading Factor	CR	AVE
Dimensi Classroom management	6	0,857	CM1	0,65	0,857	0,502
			CM2a	0,72		
			CM2b	0,79		
			CM3	0,72		
			CM4	0,72		
			CM5	0,64		
Dimensi Supportive climate	5	0,829	SC6	0,74	0,836	0,506
			SC7	0,75		
			SC8	0,74		
			SC9	0,65		
			SC10	0,67		
Dimensi Cognitive activation	7	0,755	CA11	0,39	0,755	0,324
			CA12	0,66		
			CA13	0,56		
			CA14	0,22		
			CA15	0,68		
			CA16	0,67		
			CA17	0,64		

Hasil adaptasi dan analisis *item* menghasilkan alat ukur *instructional quality* final yang terdiri dari 18 *item*. Berdasarkan hasil analisis diskriminan *item* serta pengujian reliabilitas dan validitas di atas maka dapat disimpulkan bahwa alat ukur *instructional quality* yang diadaptasikan dari Bellens et al. (2019) telah memenuhi persyaratan psikometrik yaitu memenuhi kriteria analisis *item* serta telah terbukti reliabel dan valid. Maka alat ukur *instructional quality* dapat digunakan untuk mengukur *instructional quality* dan ketiga dimensinya pada populasi siswa SMA/SMK/ sederajat di Indonesia. Berbeda dengan alat ukur *instructional quality* asli tanpa modifikasi yang mendapatkan bukti terdapat beberapa kriteria *fit* yang belum terpenuhi ($X^2/df > 3,00$; CFI < 0,90), hasil adaptasi memperoleh bukti validitas yang sudah memenuhi kriteria *fit*. Maka alat ukur hasil adaptasi disimpulkan mampu dengan akurat mengukur konstruk yang diinginkan.

Pada dimensi *classroom management* dan *supportive climate*, hasil analisis *item* menunjukkan bahwa seluruh *item* alat ukur adaptasi pada kedua konstruk tersebut signifikan dan memiliki *loading factor* yang memadai untuk menjelaskan dimensi masing-masing. Sedangkan pada dimensi *cognitive activation* ditemukan beberapa *item* yang perlu ditinjau kembali karena memiliki isu. Pada Tabel 7 terlihat bahwa *item* CA11 dan CA14 memiliki *loading factor* < 0,40. Artinya, kontribusi dari *item-item* tersebut untuk menjelaskan dimensi *cognitive activation* rendah. Pada alat ukur *instructional quality* asli yang digunakan oleh Bellens et al. (2019), hasil analisis terhadap *item* menemukan hasil serupa bahwa *item* CA11 dan CA14 memiliki *loading factor* < 0,40 atau tidak signifikan sehingga diduga bahwa pernyataan pada kedua *item* tersebut masih rendah kontribusinya untuk dapat menjelaskan dimensi yang diinginkan.

Meski begitu, pada alat ukur adaptasi ini *item* diputuskan untuk tetap dipertahankan karena melihat nilai *t-*

value yang diperoleh pada seluruh *item* sebesar $\geq 1,96$ yang menunjukkan bahwa *item-item* tersebut masih memiliki kontribusi yang cukup signifikan untuk dapat mengukur dimensi yang diinginkan. Perbedaan ditemukan pada *loading factor item* CA13 (“guru memberikan pertanyaan mengenai materi pelajaran yang harus saya pikirkan dengan sungguh-sungguh”) dimana pada alat ukur asli Bellens et al. (2019) *item* memiliki *loading factor* < 0,40 sementara pada hasil adaptasi diperoleh *loading factor* yang signifikan ($\lambda = 0,56$). Diduga hasil ini diperoleh karena adanya perbaikan pada *item* berdasarkan hasil *cognitive interview* dimana pernyataan pada *item* diberi konteks yang spesifik yaitu terkait materi pelajaran dan disampaikan secara eksplisit sehingga tidak muncul perbedaan tafsir saat alat ukur digunakan pada partisipan.

Beberapa penelitian mengenai *instructional quality* seperti penelitian Atlay et al. (2019); Fauth et al. (2019); Lazarides et al. (2021); dan Lazarides & Buchholz (2019) menggunakan metode survei kuesioner kepada siswa dalam pengumpulan data. *Item* yang digunakan pada pengukuran dipilih dan dimodifikasi sesuai dengan kebutuhan penelitian seperti penyesuaian bahasa yang digunakan serta konteks penelitian (contoh: kelas atau mata pelajaran spesifik). Analisis dari alat ukur hasil adaptasi menambahkan bukti bahwa metode kuesioner kepada siswa merupakan metode yang valid dan reliabel untuk dapat mengukur dimensi-dimensi dari *instructional quality* (Kunter & Baumert, 2006) meski eksplorasi lebih lanjut perlu dilakukan untuk dapat memperoleh alat ukur *students' rating* yang valid dan reliabel pada level kelas (Aditomo & Köhler, 2020). Selain itu, hasil proses adaptasi (*expert review*, *cognitive interview*) memberikan *insight* bahwa konteks merupakan hal yang dapat menentukan pemahaman dan pemrosesan kognitif dari partisipan terhadap makna *item*.

Sehingga alat ukur yang dirancang untuk dapat bebas digunakan dan disesuaikan dengan konteks yang diperlukan menjadi vital bagi penelitian-penelitian yang akan dilakukan selanjutnya. Diharapkan kebutuhan tersebut dapat terpenuhi oleh alat ukur hasil adaptasi yang dirancang untuk dapat digunakan secara umum di berbagai *setting* mata pelajaran/kelas tanpa perlu memodifikasi kembali *item-itemnya*. Penelitian mengenai *instructional quality* sendiri masih minim dilakukan di Indonesia. Hal ini salah satunya diduga karena belum tersedianya alat ukur *instructional quality* berbahasa Indonesia yang telah diuji cobakan validitas serta reliabilitasnya. Pada umumnya, peneliti perlu melakukan translasi terlebih dahulu dari alat ukur berbahasa asing yang digunakan pada penelitian lain (Yasin & Mustafa, 2020). Maka alat ukur hasil adaptasi telah dianalisis dan dirancang dalam Bahasa Indonesia Indonesia sehingga dapat langsung digunakan tanpa melakukan translasi kembali.

Masih terdapat beberapa keterbatasan dari penelitian ini yang dapat menjadi peluang bagi penelitian selanjutnya. Keterbatasan pertama yaitu administrasi alat ukur yang dilakukan secara daring. Keterbatasan penggunaan metode ini ialah peneliti kesulitan untuk mengontrol variabel lain yang mungkin memengaruhi keadaan diri partisipan saat mengisi kuesioner. Selain itu, keterbatasan lain dari penelitian ini yaitu hasil pengujian validitas masih menunjukkan adanya salah satu kriteria parameter yang belum fit, yaitu nilai *p-value*. Penelitian selanjutnya dapat menguji kembali reliabilitas dan validitas dari alat ukur dengan menggunakan metode lain sehingga bukti psikometrik dari alat ukur dapat menjadi semakin beragam dan lengkap. Bukti reliabilitas internal telah diperoleh sehingga selanjutnya dapat dilakukan uji seperti *test-retest* untuk mendapatkan bukti reliabilitas eksternal. Sementara untuk validitas, selain validitas konten dan konstruk yang telah diujikan dapat ditambahkan bukti validitas kriteria dimana data hasil alat ukur adaptasi dapat dibandingkan dengan alat ukur lain yang mengukur hal yang sama seperti alat ukur yang digunakan oleh Rafsanjani et al. (2020) yang mengukur *instructional quality* menggunakan kuesioner *teachers' rating*.

KESIMPULAN

Berdasarkan prosedur adaptasi dari International Test Commission (2017) Alat ukur *Instructional Quality Questionnaire for Flanders and Germany Students* dari Bellens et al. (2019) dalam bahasa Indonesia yang dikembangkan berdasarkan prosedur adaptasi dari *International Test Commission* (2017) dalam konteks umum terbukti reliabel dan valid, sehingga dapat digunakan untuk mengukur konstruk *instructional quality* guru dan dimensi-dimensinya pada siswa SMA/SMK/ sederajat di Indonesia. Alat ukur yang dihasilkan terdiri dari 18 *item* yang umum dapat digunakan pada berbagai *setting* kelas dan/atau mata pelajaran yang diajarkan guru sehingga alat ukur tidak terbatas penggunaannya pada satu *setting* tertentu. Untuk penelitian terkait pengembangan alat ukur *instructional quality* selanjutnya disarankan untuk melakukan uji coba secara tatap muka untuk dapat melakukan kontrol terhadap *setting*

pengambilan data dengan lebih baik. Kemudian disarankan pula untuk melakukan pengumpulan bukti validitas dan reliabilitas lain seperti validitas kriteria dan reliabilitas eksternal untuk menambahkan bukti psikometrik terhadap alat ukur *instructional quality*.

DAFTAR PUSTAKA

- Aditomo, A., & Köhler, C. (2020). Do student ratings provide reliable and valid information about teaching quality at the school level? *Evaluating measures of science teaching in PISA 2015*.
- Atlay, C., Tieben, N., Hillmert, S., & Fauth, B. (2019). Instructional quality and achievement inequality: How effective is teaching in closing the social achievement gap? *Learning and Instruction*, 63(April 2018), 101211. <https://doi.org/10.1016/j.learninstruc.2019.05.008>
- Azwar, S. (2009). Efek seleksi aitem berdasar daya diskriminasi terhadap reliabilitas skor tes. *Buletin Psikologi*, 17(1), 1–6.
- Baumert, J., Kunter, M., Blum, W., Klusmann, U., Krauss, S., & Neubrand, M. (2013). Cognitive activation in the mathematics classroom and professional competence of teachers: Results from the COACTIV project. *Cognitive Activation in the Mathematics Classroom and Professional Competence of Teachers: Results from the COACTIV Project*, March 2013, 1–378. <https://doi.org/10.1007/978-1-4614-5149-5>
- Bellens, K., Van Damme, J., Van Den Noortgate, W., Wendt, H., & Nilsen, T. (2019). Instructional quality: catalyst or pitfall in educational systems' aim for high achievement and equity? An answer based on multilevel SEM analyses of TIMSS 2015 data in Flanders (Belgium), Germany, and Norway. *Large-Scale Assessments in Education*, 7(1), 1–27. <https://doi.org/10.1186/s40536-019-0069-2>
- Brophy, J. (1986). Teacher influences on student achievement. *American Psychologist*, 41(10), 1069–1077. <https://doi.org/10.1037/0003-066X.41.10.1069>
- Ebel, R. L. (1967). The relation of item discrimination to test reliability. *Journal of Educational Measurement*, 4(3), 16–19.
- Fauth, B., Decristan, J., Decker, A. T., Büttner, G., Hardy, I., Klieme, E., & Kunter, M. (2019). The effects of teacher competence on student outcomes in elementary science education: The mediating role of teaching quality. *Teaching and Teacher Education*, 86, 102882. <https://doi.org/10.1016/j.tate.2019.102882>
- Fauth, B., Decristan, J., Rieser, S., Klieme, E., & Büttner, G. (2014). Student ratings of teaching quality in primary school: Dimensions and prediction of student outcomes. *Learning and Instruction*, 29, 1–9. <https://doi.org/10.1016/j.learninstruc.2013.07.001>
- Field, A. (2009). *Discovering Statistics Using SPSS* (3rd Editio). SAGE Publications Ltd. https://doi.org/10.1007/978-0-387-68969-2_13
- Förtsch, C., Werner, S., Dorfner, T., von Kotzebue, L., & Neuhaus, B. J. (2017). Effects of Cognitive Activation in Biology Lessons on Students' Situational Interest and Achievement. *Research in Science Education*, 47(3), 559–578. <https://doi.org/10.1007/s11165-016-9517-y>
- Gage, N. L., & Berliner, D. C. (1998). *Educational Psychology*. Houghton Mifflin Company.
- Guill, K., Lüdtke, O., & Köller, O. (2020). Assessing the instructional quality of private tutoring and its effects on student outcomes: Analyses from the German National Educational Panel Study. *British Journal of Educational Psychology*, 90(2), 282–300. <https://doi.org/10.1111/bjep.12281>
- Hair, J. F., Anderson, R. E., Tatham, R. L., & Black, W. C. (2019). Multivariate Data Analysis, Multivariate Data Analysis. In Book (Vol. 87, Issue 4).
- Hooper, D., Coughlan, J., & Mullen, M. R. (2008). Structural equation modelling: Guidelines for determining model fit. *Electronic Journal of Business Research Methods*, 6(1), 53–60.
- Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6(1), 1–55. <https://doi.org/10.1080/10705519909540118>
- International Test Commission. (2017). *International Test Commission Guidelines for Translating and Adapting Tests 2nd Edition*. <http://scholar.google.com/scholar?hl=en&btnG=Search&q=international+Test+Commission+Guidelines+for+Translating+and+Adapting+Tests#0>
- Kaplan, R. M., & Saccuzzo, D. P. (2013). *Psychological Testing: Principles*,

- Applications, and Issues* (J. Perkins (ed.); 8th ed.). Cengage Learning.
- Kunter, M., & Baumert, J. (2006). Who is the expert? Construct and criteria validity of student and teacher ratings of instruction. *Learning Environments Research*, 9(3), 231–251. <https://doi.org/10.1007/s10984-006-9015-7>
- Kunter, M., Frenzel, A., Nagy, G., Baumert, J., & Pekrun, R. (2011). Teacher enthusiasm: Dimensionality and context specificity. *Contemporary Educational Psychology*, 36(4), 289–301. <https://doi.org/10.1016/j.cedpsych.2011.07.001>
- Lazarides, R., & Buchholz, J. (2019). Student-perceived teaching quality : How is it related to different achievement emotions in mathematics classrooms? *Learning and Instruction*, 61(January), 45–59. <https://doi.org/10.1016/j.learninstruc.2019.01.001>
- Lazarides, R., Fauth, B., Gaspard, H., & Richard, G. (2021). *Teacher self-efficacy and enthusiasm : Relations to changes in student-perceived teaching quality at the beginning of.* 73(December 2020). <https://doi.org/10.1016/j.learninstruc.2020.101435>
- Lipowsky, F., Rakoczy, K., Pauli, C., Drollinger-Vetter, B., Klieme, E., & Reusser, K. (2009). Quality of geometry instruction and its short-term impact on students' understanding of the Pythagorean Theorem. *Learning and Instruction*, 19(6), 527–537. <https://doi.org/10.1016/j.learninstruc.2008.11.001>
- Nilsen, T., & Gustafsson, J. E. (2016). Teacher quality, instructional quality and student outcome. In *IEA Research for Education* (Vol. 2). <http://link.springer.com/10.1007/978-3-319-41252-8>
- Praetorius, A. K., Klieme, E., Herbert, B., & Pinger, P. (2018). Generic dimensions of teaching quality: the German framework of Three Basic Dimensions. *ZDM - Mathematics Education*, 50(3), 407–426. <https://doi.org/10.1007/s11858-018-0918-4>
- Praetorius, A. K., Pauli, C., Reusser, K., Rakoczy, K., & Klieme, E. (2014). One lesson is all you need? Stability of instructional quality across lessons. *Learning and Instruction*, 31, 2–12. <https://doi.org/10.1016/j.learninstruc.2013.12.002>
- Rafsanjani, M. A., Pamungkas, H. P., Ghofur, M. A., & Fitrayati, D. (2020). Exploring the key predictors of instructional quality. *ERIES Journal*, 15(4).
- Sakiz, G., Pape, S. J., & Hoy, A. W. (2012). Does perceived teacher affective support matter for middle school students in mathematics classrooms? *Journal of School Psychology*, 50(2), 235–255. <https://doi.org/10.1016/j.jsp.2011.10.005>
- Sigala, L. (2016). Hubungan Persepsi Siswa Tentang Cara Mengajar Guru Dengan Motivasi Belajar Siswa yang Pernah Mengikuti Remedial. *Psikoborneo: Jurnal Ilmiah Psikologi*, 4(3), 310–315. <https://doi.org/10.30872/psikoborneo.v4i3.4088>
- Umar, J., & Nisa, Y. F. (2020). Uji Validitas Konstruk dengan CFA dan Pelaporannya. *Jurnal Pengukuran Psikologi Dan Pendidikan Indonesia*, 9(2), 1–11.
- Willis, G. B. (2015). Analysis of Cognitive Interview in Questionnaire Design: Understanding Qualitative Research. In *Oxford University Press*. Oxford University Press, Inc. <https://ejournal.poltektegal.ac.id/index.php/siklus/article/view/298%0Ahttp://repositorio.unan.edu.ni/2986/1/5624.pdf%0Ahttp://dx.doi.org/10.1016/j.jana.2015.10.005%0Ahttp://www.biomedcentral.com/1471-2458/12/58%0Ahttp://ovidsp.ovid.com/ovidweb.cgi?T=JS&P>
- Yasin, B., & Mustafa, F. (2020). *The effect of English teacher's instructional management skills on beginner students' achievement and perception.* 13(4), 49–66.
- Yilmaz, E., Şahin, M., & Turgut, M. (2017). Variables Affecting Student Motivation Based on Academic Publications. *Journal of Education and Practice*, 8(12), 112–120. www.iiste.org
- Yu, R., & Singh, K. (2016). *Teacher support , instructional practices , student motivation , and mathematics achievement in high school.* 0671(September), 0–14. <https://doi.org/10.1080/00220671.2016.1204260>
- Yu, R., & Singh, K. (2018). Teacher support, instructional practices, student motivation, and mathematics achievement in high school. *Journal of Educational Research*, 111(1), 81–94. <https://doi.org/10.1080/00220671.2016.1204260>

LAMPIRAN

Lampiran 1. Alat Ukur *Instructional Quality* Hasil Adaptasi

Instruksi:

Bacalah pernyataan-pernyataan yang tertera, kemudian pilihlah salah satu jawaban yang menunjukkan keadaan kelas yang Anda ikuti. Terdapat 4 pilihan jawaban yaitu sangat tidak setuju (1), tidak setuju (2), setuju (3), dan sangat setuju (4). Setiap pernyataan harus diisi dengan 1 jawaban saja, dan jangan sampai ada yang terlewat.

Kuesioner ini merujuk kepada keadaan kelas (nama mata pelajaran/kelas yang hendak diteliti/diukur, contoh: kelas Bahasa Indonesia) yang Anda ikuti saat ini, pada mata pelajaran yang diajar oleh (guru mata pelajaran/kelas yang hendak diteliti/diukur, contoh: guru Bahasa Indonesia). Sebelum menjawab pertanyaan-pertanyaan di bawah, ingat-ingat terlebih dahulu situasi belajar yang biasa Anda alami selama proses belajar daring (online) Anda bersama guru tersebut. Seberapa jauh Anda menyetujui pernyataan-pernyataan di bawah ini?

Kode Item	Pernyataan
CM1.	Guru mampu membuat siswa mendengarkan materi pelajaran yang sedang disampaikan
CM2a.	Guru mampu menenangkan suasana kelas yang berisik
CM2b.	Guru mampu membuat siswa fokus terhadap kegiatan belajar di kelas sehingga siswa tidak melakukan aktivitas lain yang mengganggu kegiatan belajar
CM3.	Guru mampu dengan cepat membuat siswa tenang sebelum memulai pelajaran
CM4.	Guru mampu membuat siswa mematuhi peraturan kelas yang telah ditentukan (misal, tidak mengobrol sendiri di kelas)
CM5.	Guru mampu mengarahkan siswa untuk menggunakan waktu belajar di kelas secara efektif
SC6.	Guru memberi dukungan kepada saya bahkan ketika saya membuat kesalahan saat mengerjakan latihan/tugas
SC7.	Guru menunjukkan sikap peduli kepada saya
SC8.	Guru memberitahu saya cara agar menjadi lebih baik ketika membuat kesalahan saat mengerjakan latihan/tugas
SC9.	Guru bersikap baik kepada saya
SC10.	Guru saya menunjukkan rasa percaya bahwa saya dapat menyelesaikan tugas yang sulit
CA11.	Guru memberi tugas yang membuat saya perlu berpikir dengan sungguh-sungguh
CA12.	Guru menanyakan materi apa yang sudah saya pahami dan apa yang belum
CA13.	Guru memberikan pertanyaan mengenai materi pelajaran yang harus saya pikirkan dengan sungguh-sungguh
CA14.	Guru memberi tugas yang pada awalnya terlihat sulit
CA15.	Guru menanyakan hal apa yang telah saya ketahui tentang materi yang baru
CA16.	Guru memberi tugas yang membuat saya termotivasi untuk menyelesaikannya walaupun sulit
CA17.	Guru meminta saya untuk dapat menjelaskan jawaban dari tugas yang saya kerjakan