



Jurnal

Penelitian dan Evaluasi Pendidikan

- Lukas Lui Uran* Evaluasi Implementasi KTSP dan Kurikulum 2013 pada SMK se-Kabupaten Belu, Nusa Tenggara Timur
- Akbar Iskandar, Mubammad Rizal* Analisis Kualitas Soal di Perguruan Tinggi Berbasis Aplikasi TAP
- Weni Wendari, Samsul Hadi* Estimasi Kesalahan Pengukuran *Standard Setting* dalam Penilaian Kompetensi Matematika Tingkat SMP di Kabupaten Sumbawa
- Riyan Arthur* Evaluasi Program Diklat Karya Tulis Ilmiah untuk Widyaiswara Pusbangtendik Kemdikbud
- Susilowati, Sajidan, Murni Ramli* Keefektifan Perangkat Pembelajaran Berbasis *Inquiry Lesson* untuk Meningkatkan Keterampilan Berpikir Kritis Siswa
- Lantip Diat Prasajo, Fredrik Abia Kande, Amirul Mukminin* Evaluasi Pelaksanaan Standar Proses Pendidikan pada SMP Negeri di Kabupaten Sleman
- Lilik Sabdaningtyas* Model Evaluasi Implementasi Kebijakan Standar Pelayanan Minimal pada Satuan Pendidikan Jenjang SD
- Wahyu Hidayat, Jamil Bin Ahmad, Mohd. Isa Bin Hamzah* Nilai Keutamaan Pengetahuan dan Kebijaksanaan dalam Konteks Pendidikan Karakter Bangsa
- Wirman Kasmayadi, Kumaidi, Sumarno* Pengembangan Instrumen Asesmen Diri Guru terhadap Aktivitas Pengembangan Keprofesian Berkelanjutan
- Marjuki, Djemari Mardapi, Badrun Kartowagiran* Pengembangan Model Akreditasi Sekolah Menengah Atas/Madrasah Aliyah (SMA/MA)



Jurnal Penelitian dan Evaluasi Pendidikan

Volume 22, No 1, June 2018



Jurnal

Penelitian dan Evaluasi Pendidikan

HIMPUNAN EVALUASI PENDIDIKAN INDONESIA (HEPI)
in cooperation with
GRADUATE SCHOOL OF UNIVERSITAS NEGERI YOGYAKARTA
Kampus Karangmalang, Yogyakarta 55281. Phone. 0274 550836 Fax : 0274 520326

Website: <http://journal.uny.ac.id/index.php/jpep>
e-mail: jurnalhepi@uny.ac.id



9 772338 606001



1410 4725



Jurnal

Penelitian dan Evaluasi Pendidikan

Publisher

HIMPUNAN EVALUASI PENDIDIKAN INDONESIA
In Cooperation With
PROGRAM PASCASARJANA UNIVERSITAS NEGERI YOGYAKARTA
(MOU Nomor 195 B/J.35.17/LK/04)

Director of Publication

Djemari Mardapi, *Universitas Negeri Yogyakarta*

Editor in Chief

Samsul Hadi, *Faculty of Engineering, Universitas Negeri Yogyakarta*

Associate Editors

Nur Hidayanto Pancoro Setyo Putro, *Faculty of Languages & Art, Universitas Negeri Yogyakarta*

Editors

Edi Istiyono, *Faculty of Mathematics and Natural Sciences, Universitas Negeri Yogyakarta*

Badrun Kartowagiran, *Faculty of Engineering, Universitas Negeri Yogyakarta*

Sudiyatno, *Faculty of Engineering, Universitas Negeri Yogyakarta*

Jailani, *Faculty of Mathematics and Natural Sciences, Universitas Negeri Yogyakarta*

Alita Arifiana Anisa, *Lembaga Pendidikan dan Pengembangan Profesi Indonesia*

Board of Reviewers

Suratno

Universitas Lambung Mangkurat

Moch Alip

Universita Negeri Yogyakarta

Budiono

Universitas Sebelas Maret Surakarta

Maria Josephine Wantah

Universitas Negeri Manado

Kartono

Universitas Negeri Semarang

Rasmuin

Universitas Dayanu Ikhsanuddin Bau-Bau

Purwo Susongko

Universitas Pancasakti Tegal

Nurul Choyimah

LAIN Tullungagung

Zurqoni

STAIN Samarinda

Amir Syamsudin

Universitas Negeri Yogyakarta

Wasis

Universitas Negeri Surabaya

Sugeng Eko Putro Widoyoko

Universitas Muhammadiyah Purworejo

Nanik Estidarsani

Universitas Negeri Surabaya

Ekohariadi

Universitas Negeri Surabaya

Mansyur

Universitas Negeri Makassar

Undang Rosidin

Universitas Lampung

Syukrul Hamdi

Universitas Hamzanwadi

Lilik Sabdaningtyas

Universitas Lampung

Primardiana Hermilia Wijayati

Universitas Negeri Malang

Correspondence: Graduate School of Universitas Negeri Yogyakarta
Kampus Karangmalang, Yogyakarta, 55281, Telp. (0274) 550835, Fax. (0274) 520326

Homepage: <http://journal.uny.ac.id/index.php/jpep>

e-mail: jurnalhepi@uny.ac.id

FOREWORDS

We are very pleased that *Jurnal Penelitian dan Evaluasi Pendidikan* is releasing its issue **Volume 22, No 1, June 2018**. We are also very excited that the journal has been attracting papers from many institutions in Indonesia. *Jurnal Penelitian dan Evaluasi Pendidikan* was first published in **1998** and since then regularly published online and in print twice a year: June and December.

Jurnal Penelitian dan Evaluasi Pendidikan with ISSN 1410-4725 (*printed*) and ISSN 2338-6061 (*online*) has been **re-accredited** by Indonesian Ministry of Education and Culture decision Number 040/P/2014 which is valid for 5 (five) years since enacted on 18 Februari 2014.

Jurnal Penelitian dan Evaluasi Pendidikan is a showcase of original, rigorously conducted educational evaluation, measurement and assessment from primary, secondary, and higher education institutions. Each issue of this journal is not limited to comprehensive syntheses of studies towards developing new understandings of educational evaluation, measurement and assessment only, but also explores scholarly analyses of issues and trends in the field.

Yogyakarta, June 2018

Editor in Chief

Table of Content

<i>Lukas Lwi Uran</i>	Evaluasi Implementasi KTSP dan Kurikulum 2013 pada SMK se-Kabupaten Belu, Nusa Tenggara Timur	1-11
<i>Akbar Iskandar, Muhammad Rizal</i>	Analisis Kualitas Soal di Perguruan Tinggi Berbasis Aplikasi TAP	12-23
<i>Weni Wendari, Samsul Hadi</i>	Estimasi Kesalahan Pengukuran <i>Standard Setting</i> dalam Penilaian Kompetensi Matematika Tingkat SMP di Kabupaten Sumbawa	24-34
<i>Riyan Arthur</i>	Evaluasi Program Diklat Karya Tulis Ilmiah untuk Widyaiswara Pusbangtendik Kemdikbud	35-48
<i>Susilowati, Sajidan, Murni Ramli</i>	Keefektifan Perangkat Pembelajaran Berbasis <i>Inquiry Lesson</i> untuk Meningkatkan Keterampilan Berpikir Kritis Siswa	49-60
<i>Lantip Diat Prasajo, Fredrik Abia Kande, Amirul Mukeminin</i>	Evaluasi Pelaksanaan Standar Proses Pendidikan pada SMP Negeri di Kabupaten Sleman	61-69
<i>Lilik Sabdaningtyas</i>	Model Evaluasi Implementasi Kebijakan Standar Pelayanan Minimal pada Satuan Pendidikan Jenjang SD	70-82
<i>Wahyu Hidayat, Jamil Bin Ahmad, Mohd. Isa Bin Hamzah</i>	Nilai Keutamaan Pengetahuan dan Kebijaksanaan dalam Konteks Pendidikan Karakter Bangsa	83-91
<i>Wirman Kasmayadi, Kumaidi, Sumarno</i>	Pengembangan Instrumen Asesmen Diri Guru terhadap Aktivitas Pengembangan Keprofesian Berkelanjutan	92-104
<i>Marjuki, Djemari Mardapi, Badrun Kartowagiran</i>	Pengembangan Model Akreditasi Sekolah Menengah Atas /Madrasah Aliyah (SMA/MA)	105-117

EVALUASI IMPLEMENTASI KTSP DAN KURIKULUM 2013 PADA SMK SE-KABUPATEN BELU, NUSA TENGGARA TIMUR

Lukas Lui Uran

SMK St. Yosef Nenuk

Atambua, Naekasa, Kec. Tasifeto Barat, Kab. Belu, Nusa Tenggara Timur, Indonesia

Email: lukas_lui@yahoo.co.id

Abstrak

Tujuan penelitian ini adalah mengungkapkan ketercapaian pelaksanaan Kurikulum Tingkat Satuan Pendidikan (KTSP) dan Kurikulum 2013, di SMK se-Kabupaten Belu yang meliputi: (1) ketercapaian pelaksanaan pembelajaran yang dilihat dari konteks, input, proses, dan produk, (2) hambatan yang dialami guru dalam pelaksanaan pembelajaran, dan (3) kelebihan dan kekurangan pembelajaran dengan menggunakan kedua kurikulum. Penelitian ini merupakan penelitian evaluasi dengan pendekatan deskriptif kuantitatif dengan model CIPP. Subjek penelitian ini adalah guru dan siswa SMK yang diambil dengan menggunakan teknik *probability sampling*. Hasil penelitian menunjukkan: (1) hasil evaluasi implementasi KTSP dan Kurikulum 2013 dilihat dari segi konteks, input proses dan produk termasuk dalam kategori sangat berhasil; (2) hambatan utama yang dialami guru dalam pelaksanaan pembelajaran dengan menggunakan KTSP adalah keterbatasan literatur sedangkan hambatan utama dalam Kurikulum 2013 adalah kesulitan dalam membuat format penilaian pembelajaran; (3) pada KTSP guru fleksibel dalam mengembangkan kurikulum, namun mengalami keterbatasan literatur. Kurikulum 2013 mampu meningkatkan kreativitas guru dan siswa, namun sulit dipahami.

Kata kunci: *evaluasi, implementasi, CIPP*

AN EVALUATION OF THE IMPLEMENTATION OF KTSP & CURRICULUM 2013 AT VOCATIONAL HIGH SCHOOL IN BELU REGENCY, EAST NUSA TENGGARA

Abstract

The aim of this research was to reveal the achievement level of *Kurikulum Tingkat Satuan Pendidikan* (KTSP) and Curriculum 2013 implementation at Vocational High School (SMK) in Belu, including: (1) the achievement level of teaching processes based on context, input, process, and output; (2) teacher's constraints in the implementation of teaching processes; and (3) the strengths and the weaknesses of teaching process of the two curriculums. This study was an evaluation research using quantitative descriptive approach and CIPP model. The subjects of this research were teachers and students of vocational high school in Belu, East Nusa Tenggara, chosen by using probability sampling technique. The results of this study show that: (1) the result of evaluation on KTSP and Curriculum 2013 implementation in terms of context, input, process and output is considered as highly successful; (2) the teachers' main constraint in teaching by using KTSP is the lack of literature, while in teaching by using Curriculum 2013, their main constraint is the lack of proper skills in creating the learning assessment form; (3) By using KTSP, teachers have flexibility in developing the curriculum but they have to face the lack of the literature. Whereas, the implementation of Curriculum 2013 can improve teachers' and students' creativity, but it is difficult to be understood.

Keywords: *evaluation, implementation, CIPP*

Permalink/DOI: <http://dx.doi.org/10.21831/pep.v22i1.13309>

Pendahuluan

Pendidikan seperti yang dijelaskan dalam UU No. 20 Tahun 2003 tentang Sistem Pendidikan Nasional (Sisdiknas) (Depdiknas, 2003) adalah usaha sadar dan terencana untuk mewujudkan suasana belajar dan proses pembelajaran agar peserta didik secara aktif mengembangkan potensi dirinya untuk memiliki kekuatan spriritual keagamaan, pengendalian diri, kepribadian, kecerdasan, akhlak mulia, serta keterampilan yang diperlukan dirinya, masyarakat bangsa dan negara. Dari definisi pendidikan di atas, terkandung makna dan tujuan yang sangat penting dan mulia, mencakup seluruh aspek untuk memanusiakan manusia. Guna mencapai tujuan pendidikan tersebut dibutuhkan usaha dan dukungan berbagai komponen pendidikan yang saling berkaitan dan saling mempengaruhi. Salah satu di antaranya adalah kurikulum pendidikan.

Kurikulum pendidikan adalah seperangkat rencana dan pengaturan mengenai tujuan, isi, dan bahan pelajaran serta cara yang digunakan sebagai pedoman penyelenggaraan kegiatan pembelajaran untuk mencapai tujuan pendidikan tertentu. Pengertian ini termuat dalam Undang-Undang No. 20 Tahun 2003 tentang Sistem Pendidikan Nasional. Dalam kurikulum terkandung daftar mata pelajaran, rencana pembelajaran, hasil terstruktur pembelajaran dan sebagainya. Singkatnya kurikulum adalah pedoman yang mengatur bagaimana pendidikan dijalankan untuk mencapai suatu tujuan pendidikan.

Kurikulum sebagai pedoman yang mengatur pendidikan di Indonesia telah mengalami beberapa kali pergantian. Pergantian kurikulum ini terjadi karena kurikulum tersebut dinilai sudah tidak relevan lagi dengan realitas, perubahan dan tantangan dunia pendidikan. Indonesia telah mengalami sembilan kali pergantian kurikulum sejak tahun 1947 sampai dengan tahun 2013. Dua kurikulum terakhir yang berlaku adalah Kurikulum 2006 atau Kurikulum Tingkat Satuan Pendidikan (KTSP). Kurikulum ini merupakan pengganti kurikulum 2004 atau Kurikulum Berbasis Kompetensi

(KBK). Sedangkan kurikulum yang lainnya adalah Kurikulum 2013 atau biasa disebut K.13. Kurikulum 2013 merupakan pengganti dari KTSP.

Kurikulum Tingkat Satuan Pendidikan (KTSP) merupakan kurikulum yang dirancang oleh pemerintah sebagai upaya mencapai keunggulan masyarakat dalam penguasaan ilmu dan teknologi seperti yang digariskan dalam haluan negara. KTSP juga merupakan suatu konsep yang menawarkan otonomi bagi sekolah untuk menentukan kebijakannya dengan tujuan untuk meningkatkan mutu dan efisiensi pendidikan (Susilo, 2007, p. 12). Lebih lanjut Mulyasa (2006, p. 20), menjelaskan KTSP adalah kurikulum operasional yang disusun, dikembangkan dan dilaksanakan oleh satuan pendidikan yang sudah siap dan mampu mengembangkannya dengan memperhatikan UU No. 20 Tahun 2003 tentang Sistem Pendidikan Nasional. Tujuan utama KTSP adalah memandirikan dan memberdayakan sekolah dalam mengembangkan kompetensi yang akan disampaikan kepada peserta didik sesuai dengan kondisi lingkungan (Susilo, 2007, p. 13). Sebagai kurikulum pendidikan, KTSP memiliki sejumlah karakteristik. Menurut Muslich (2007, p. 11) karakteristik KTSP adalah: (1) berpusat pada potensi, perkembangan, kebutuhan, dan kepentingan peserta didik dan lingkungannya, (2) beragam dan terpadu, (3) tanggap terhadap perkembangan iptek dan seni, (4) relevan dengan kebutuhan kehidupan, (5) menyeluruh dan berkesinambungan, (6) belajar sepanjang hayat dan (7) seimbang antara kepentingan nasional dan kepentingan daerah.

Berdasarkan telaah tersebut, dapat disimpulkan bahwa KTSP adalah kurikulum yang memberikan otonomi kepada penyelenggara pendidikan (sekolah) untuk menyelenggarakan pembelajaran sesuai dengan kondisi sekolah masing-masing. Tujuannya adalah meningkatkan mutu dan efisiensi pendidikan. KTSP resmi diimplementasikan pada tahun 2007.

Enam tahun penyelenggaraan pendidikan dengan menggunakan KTSP sebagai kurikulum nasional, Pemerintah merasa per-

lu untuk membuat pembaharuan dan perubahan kurikulum. Oleh sebab itu melalui Kementerian Pendidikan dan Kebudayaan Indonesia, Pemerintah menyiapkan dan menetapkan kurikulum baru yang disebut dengan Kurikulum 2013. Sejak pertama kali konsepnya diperkenalkan, kurikulum ini mendapat banyak perhatian dan tanggapan pro dan kontra dari berbagai kalangan. Meskipun demikian, pemerintah mengambil sikap untuk tetap mengimplementasikan Kurikulum 2013 pada tahun ajaran baru bulan Juli 2013.

Kurikulum 2013 mempunyai beberapa karakteristik. Secara umum Kurikulum 2013 mempunyai karakteristik sebagai berikut: (a) belajar tuntas, yaitu peserta didik tidak diperkenankan mengerjakan pekerjaan berikutnya sebelum mampu menyelesaikan pekerjaan dengan prosedur yang benar, (b) penilaian autentik, (c) penilaian berkesinambungan, penilaian dilakukan secara terus menerus dan berkesinambungan selama pembelajaran berlangsung, (d) menggunakan teknik penilaian yang bervariasi. Teknik penilaian yang dipilih dapat berupa tertulis, lisan, produk, portofolio, unjuk kerja, proyek, pengamatan, dan penilaian diri, dan (e) berdasarkan acuan kriteria.

Selain karakteristik umum tersebut, menurut Peraturan Menteri No. 70 tahun 2013 (Menteri Pendidikan dan Kebudayaan Republik Indonesia, 2013), Kurikulum 2013 memiliki karakteristik lain yaitu (1) Isi atau konten kurikulum yaitu kompetensi dinyatakan dalam bentuk Kompetensi Inti (KI) kelas dan dirinci lebih lanjut dalam Kompetensi Dasar (KD) mata pelajaran. (2) Kompetensi Inti (KI) merupakan gambaran secara kategorial mengenai kompetensi sikap, pengetahuan, dan keterampilan (kognitif dan psikomotor). (3) Kompetensi Dasar (KD) merupakan kompetensi yang dipelajari peserta didik untuk mata pelajaran di kelas tertentu. (4) Kompetensi Inti dan Kompetensi Dasar di jenjang pendidikan dasar diutamakan pada ranah sikap sedangkan pada jenjang pendidikan menengah pada kemampuan intelektual (kemampuan kognitif tinggi). (5) Kompetensi Inti men-

jadi unsur organisatoris (*organizing elements*). (6) Kompetensi Dasar yang dikembangkan didasarkan pada prinsip akumulatif, saling memperkuat (*reinforced*) dan memperkaya (*enriched*). (7) Silabus dikembangkan sebagai rancangan belajar untuk satu tema satu mata pelajaran (SMP/MTS, SMA/MA, SMK/MAK). (8) Rencana Pelaksanaan Pembelajaran dikembangkan dari setiap KD.

KTSP maupun Kurikulum 2013 memiliki perbedaan, mempunyai kelebihan dan kekurangan masing-masing dan merupakan sebuah produk kebijakan dari pemerintah dalam sektor pendidikan. Namun pelaksanaan di lapangan seringkali tidak dievaluasi. Padahal evaluasi sangat bermanfaat untuk mengumpulkan, menilai dan mengambil keputusan terhadap implementasi dan efektivitas suatu program termasuk kurikulum. Stufflebeam & Shinkfield (2014, p. 7) menyatakan bahwa, evaluasi merupakan suatu proses menyediakan informasi yang dapat dijadikan sebagai pertimbangan untuk menentukan harga dan jasa (*the worth and merit*) dari tujuan yang dicapai, desain, implementasi dan dampak untuk membantu pembuatan keputusan, pertanggung jawaban dan peningkatan pemahaman terhadap fenomena.

Berdasarkan pengertian evaluasi tersebut, dapat ditarik kesimpulan bahwa evaluasi kurikulum memegang peranan penting baik dalam penentuan kebijaksanaan pendidikan, maupun pada pengambilan keputusan dalam kurikulum. Hasil-hasil evaluasi kurikulum dapat digunakan oleh para pemegang kebijakan pendidikan dan para pengembang kurikulum dalam memilih dan menetapkan kebijaksanaan pengembangan sistem pendidikan dan model pengembangan kurikulum yang digunakan. Hasil evaluasi kurikulum juga dapat dipakai oleh guru, kepala sekolah maupun para pelaksana pendidikan lainnya untuk mengetahui perkembangan siswa, memilih bahan pelajaran, metode serta teknik penilaian pendidikan. Evaluasi kurikulum memerlukan model evaluasi yang tepat agar dapat mengarah pada perbaikan. Salah satu model evaluasi tersebut adalah CIPP.

Model CIPP merupakan kepanjangan dari *Context, Input, Process*, dan *Product*. Model ini dipandang tepat untuk mengevaluasi kurikulum karena penilaiannya bersifat menyeluruh menyangkut aspek-aspek penting kurikulum dan mengarah kepada perbaikan. Hal ini sejalan dengan konsep yang ditawarkan oleh (Daniel L. Stufflebeam, Madaus, & Kellaghan (2002, p. 280) yaitu, “*The CIPP approach is based on the view that the most important purpose of evaluation is not to prove but to improve*” yang bermakna model evaluasi CIPP bertujuan bukan untuk membuktikan melainkan untuk memperbaiki.

Kabupaten Belu, Provinsi Nusa Tenggara Timur salah satu kabupaten yang juga mengalami dinamika perubahan kurikulum baik KTSP maupun Kurikulum 2013. Sebelumnya, pembelajaran di Kabupaten ini menggunakan KTSP, namun ketika Kurikulum 2013 ditetapkan sebagai kurikulum nasional, pembelajaran di kabupaten ini menggunakan Kurikulum 2013, Implementasi Kurikulum 2013 di Kabupaten Belu tidak berlangsung lama. Pembelajaran dengan menggunakan Kurikulum 2013 hanya berjalan selama tiga semester bahkan terdapat sekolah yang hanya menggunakan Kurikulum 2013 selama dua semester pembelajaran. Selanjutnya, sekolah memakai kembali Kurikulum KTSP sebagai pedoman dalam kegiatan pembelajaran.

Proses perubahan kurikulum yang terbilang cepat memunculkan beragam respon dan hasil. Oleh karena itu, penelitian bertujuan untuk meneliti bagaimana implementasi KTSP maupun Kurikulum 2013 pada sekolah menengah kejuruan di Kabupaten Belu. Penelitian ini mencakup kekurangan dan kelebihan kedua kurikulum dan kendala-kendala yang dihadapi dalam pelaksanaan kedua kurikulum dan bagaimana upaya untuk mengatasi kendala-kendala tersebut.

Metode

Penelitian ini dikategorikan sebagai penelitian evaluatif (*Evaluation Research*) dengan menggunakan model CIPP. Penelitian ini bertujuan untuk menilai konteks, input, proses dan produk. Selain itu, penelitian ini

juga mencakup hambatan, kekurangan dan kelebihan dari KTSP maupun Kurikulum 2013. Penelitian evaluasi ini dilaksanakan di SMK se-Kabupaten Belu, Nusa Tenggara Timur. Pelaksanaan penelitian meliputi tiga tahap, yaitu tahap prasurvei, tahap analisis data dan penulisan laporan akhir.

Subjek evaluasi dalam penelitian ini adalah pihak Dinas Pendidikan Kabupaten Belu, kepala sekolah, guru dan siswa pada SMK se-Kabupaten Belu, Nusa Tenggara Timur yang pernah melaksanakan pembelajaran dengan menggunakan KTSP dan Kurikulum 2013. SMK di Kabupaten Belu berjumlah 8 sekolah yang terdiri dari 5 SMK negeri dan 3 SMK swasta namun hanya 5 SMK yang pernah melaksanakan KTSP dan Kurikulum 2013 yaitu SMK Negeri 1 Atambua, SMK Negeri 1 Belu, SMK Kakuluk Mesak, SMK Kusuma dan SMK St. Yosef Nenuk. Secara rinci, responden yang menjadi sampel dalam penelitian ini adalah 5 (lima) Kepala sekolah, 70 dari 285 guru dan 110 siswa.

Data yang telah terkumpul selanjutnya dianalisis dengan menggunakan analisis deskriptif kuantitatif. Data dari kuesioner/angket dianalisis secara kuantitatif sedangkan data hasil wawancara dan dokumentasi akan melengkapi data tersebut. Hasil analisis pelaksanaan kurikulum akan dibandingkan dengan kriteria yang ditentukan sebagaimana yang tampak pada Tabel 1.

Tabel 1. Kategori Skor Penelitian

No.	Skor	Kategori
1.	$X \geq \bar{X} + 1.SBx$	Sangat berhasil
2.	$X + 1.SBx > X \geq \bar{X}$	Berhasil
3.	$\bar{X} > X \geq \bar{X} - 1.SBx$	Kurang berhasil
4.	$X < \bar{X} - 1.SBx$	Tidak Berhasil

Sumber: Mardapi (2012, p. 162)

Keterangan

X = skor

\bar{X} = rata-rata hitung

SBx = simpangan baku

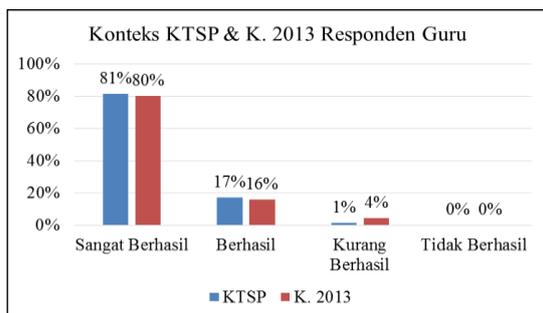
Hasil dan Pembahasan

Berdasarkan penelitian, dapat diketahui beberapa informasi mengenai implemen-

tasi KTSP dan Kurikulum 2013 pada pembelajaran di SMK se-Kabupaten Belu NTT yang dilihat berdasarkan kriteria *context, input, process, product*.

Evaluasi Konteks KTSP dan Kurikulum 2013

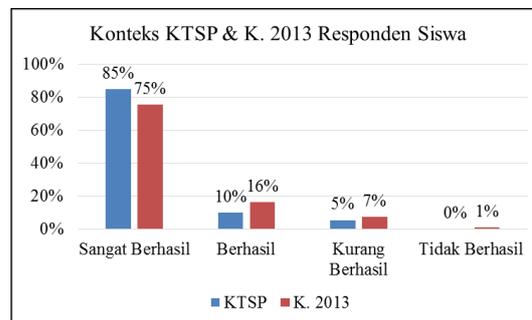
Evaluasi konteks diperoleh dari menganalisis tujuan, manfaat dan sasaran implementasi KTSP dan Kurikulum 2013. Hasil penilaian komponen konteks dengan responden guru pada KTSP dan Kurikulum 2013 ditunjukkan melalui Gambar 1.



Gambar 1. Penyebaran Jawaban Konteks KTSP & Kurikulum 2013 (responden guru)

Hasil tersebut menunjukkan bahwa implementasi KTSP dan Kurikulum 2013 pada aspek konteks masuk dalam kategori sangat berhasil dengan persentase yang berbeda. KTSP memperoleh persentase 81% dan Kurikulum 2013 sebesar 80%. Keberhasilan tertinggi baik KTSP dan Kurikulum 2013 terdapat pada indikator pemahaman para guru terhadap tujuan, manfaat dan sasaran yang dicapai setiap kurikulum. Pada kategori berhasil KTSP memperoleh persentase lebih besar daripada Kurikulum 2013 dengan persentase 17% untuk KTSP dan 16% untuk Kurikulum 2013. Pada kategori kurang berhasil, Kurikulum 2013 memperoleh persentase lebih besar daripada KTSP dengan persentase 4% untuk Kurikulum 2013 dan 1% untuk KTSP. Kekurangberhasilan kedua kurikulum ini terdapat pada indikator mendorong terwujudnya otonomi sekolah dan kerja sama dengan forum MGMP/KKG. Sedangkan untuk kategori tidak berhasil kedua kurikulum memperoleh persentase sebesar 0%.

Selanjutnya, hasil analisis evaluasi konteks berdasarkan kuesioner yang dibagikan kepada siswa dapat dilihat pada Gambar 2.



Gambar 2. Penyebaran Jawaban Konteks KTSP (responden siswa)

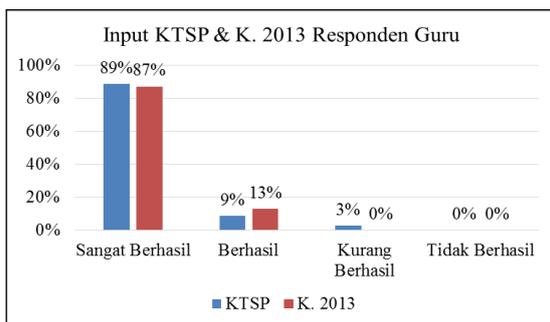
Hasil tersebut menunjukkan bahwa implementasi KTSP dan Kurikulum 2013 pada aspek konteks termasuk dalam kategori sangat berhasil. KTSP memperoleh persentase lebih besar daripada Kurikulum 2013 dengan persentase 85% untuk KTSP dan 75% untuk Kurikulum 2013. Keberhasilan yang tertinggi dicapai pada indikator kemandirian siswa dalam belajar dan memberikan kesempatan yang luas kepada siswa untuk mempergunakan sarana. Pada kategori berhasil KTSP memperoleh persentase lebih kecil daripada Kurikulum 2013 dengan persentase 10% untuk KTSP dan 16% untuk Kurikulum 2013. Pada kategori kurang berhasil Kurikulum 2013 memperoleh persentase lebih besar dari pada KTSP dengan persentase 7% untuk Kurikulum 2013 dan 5% untuk KTSP. Kekurangberhasilan ini terdapat pada indikator inisiatif siswa dalam mengembangkan potensi yang dimilikinya. Sedangkan untuk kategori tidak berhasil Kurikulum 2013 memperoleh persentase sebesar 1% sedangkan KTSP sebesar 0%.

Berdasarkan hasil penelitian ini, baik untuk responden guru maupun siswa dari aspek konteks, implementasi KTSP dan Kurikulum 2013 termasuk dalam kategori sangat berhasil. Sehingga dapat disimpulkan bahwa aspek konteks sangat mendukung pelaksanaan KTSP dan Kurikulum 2013 pada SMK se-Kabupaten Belu. Hal ini terlihat dari indikator keberhasilan implementasi yaitu adanya visi dan misi sekolah, tujuan program yang disusun dan prediksi ke-

butuhan ke depan, juga terlaksananya penerimaan siswa setiap tahunnya pada setiap SMK ini.

Evaluasi *Input* KTSP dan Kurikulum 2013

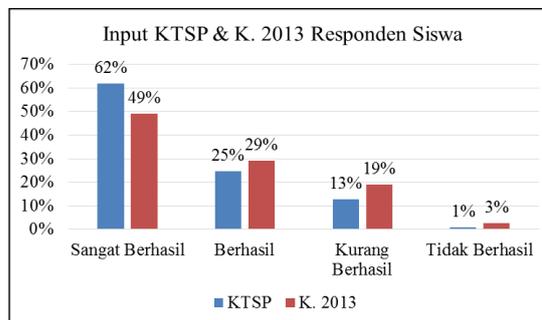
Evaluasi *input* diperoleh dari pengukuran sejauh mana kesiapan sekolah, kondisi sarana prasarana, sosialisasi, ketersediaan dana, dan buku-buku. Hasil penilaian komponen *input* dengan responden guru pada KTSP dan Kurikulum 2013 ditunjukkan dalam Gambar 3.



Gambar 3. Penyebaran Jawaban Input KTSP & Kurikulum 2013 (responden guru)

Hasil tersebut menunjukkan bahwa implementasi KTSP dan Kurikulum 2013 untuk aspek *input* termasuk dalam kategori sangat berhasil. KTSP memperoleh persentase lebih besar daripada Kurikulum 2013 dengan persentase 89% untuk KTSP dan 87% untuk Kurikulum 2013. Keberhasilan tertinggi diperoleh pada indikator kesediaan sumber daya manusia dan dana biaya operasional sekolah. Pada kategori berhasil, KTSP memperoleh persentase lebih kecil daripada Kurikulum 2013 dengan persentase 9% untuk KTSP dan 13% untuk Kurikulum 2013. Pada kategori kurang berhasil KTSP memperoleh persentase lebih besar yaitu 3% dan untuk Kurikulum 2013 sebesar 0% atau tidak ada responden yang menyatakan kurang berhasil. Kekurangberhasilan terdapat pada indikator sosialisasi kurikulum dan tersedianya buku-buku. Sedangkan untuk kategori tidak berhasil kedua kurikulum memperoleh persentase sebesar 0% atau tidak ada responden yang menyatakan bahwa kedua kurikulum tidak berhasil.

Selanjutnya, hasil analisis evaluasi input berdasarkan kuesioner yang dibagikan kepada siswa dapat dilihat pada Gambar 4.



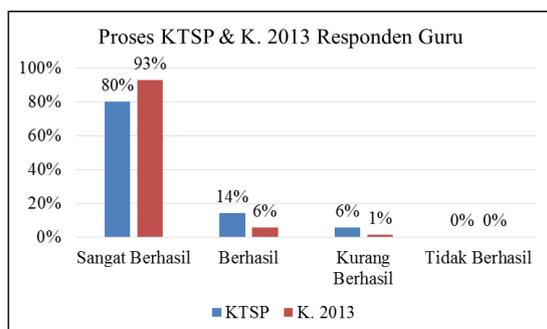
Gambar 4. Penyebaran Jawaban Input KTSP & Kurikulum 2013 (responden siswa)

Hasil tersebut menunjukkan bahwa implementasi KTSP dan Kurikulum 2013 pada aspek *input* termasuk dalam kategori sangat berhasil. KTSP memperoleh persentase sebesar 62% dan Kurikulum 2013 sebesar 49%. Indikator keberhasilan aspek ini adalah guru yang berkompeten Pada kategori berhasil KTSP memperoleh persentase sebesar 25% dan Kurikulum 2013 sebesar 29% Selanjutnya pada kategori kurang berhasil KTSP memperoleh persentase sebesar 13% dan Kurikulum 2013 sebesar 19%. Ketidakterhasilan terdapat pada indikator membutuhkan buku-buku sumber yang memadai dan memberikan akses kepada siswa untuk menggunakan sarana sekolah di luar jam sekolah. Pada kategori tidak berhasil KTSP memperoleh persentase lebih kecil yakni sebesar 1% sedangkan Kurikulum 2013 sebesar 3%.

Hasil penelitian ini menunjukkan bahwa implementasi KTSP dan Kurikulum 2013 pada aspek *input* termasuk dalam kategori sangat berhasil untuk responden guru dan siswa. Itu berarti bahwa sekolah siap menerapkan kurikulum. Kesiapan tersebut meliputi tersedianya tenaga pengajar, sarana prasarana, sosialisasi pelaksanaan kurikulum, dana, dan buku-buku yang diperlukan untuk menunjang proses pembelajaran.

Evaluasi Proses KTSP dan Kurikulum 2013

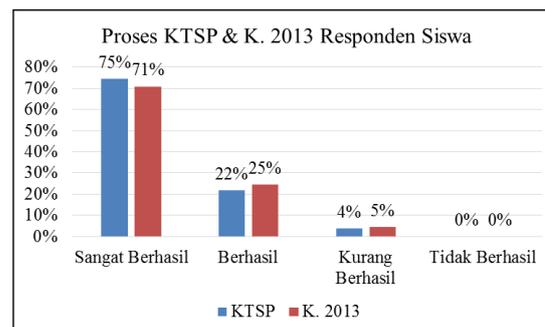
Evaluasi proses diperoleh dari analisis terhadap proses implementasi KTSP dan Kurikulum 2013. Hasil penilaian komponen proses dengan responden guru pada KTSP dan Kurikulum 2013 ditunjukkan melalui Gambar 5.



Gambar 5. Penyebaran Jawaban Proses KTSP & Kurikulum 2013 (responden guru)

Hasil tersebut menunjukkan bahwa implementasi KTSP dan Kurikulum 2013 pada aspek proses termasuk dalam kategori sangat berhasil. KTSP memperoleh persentase lebih kecil daripada Kurikulum 2013 yakni sebesar 80% untuk KTSP dan 93% untuk Kurikulum 2013. Keberhasilan aspek proses terdapat pada indikator pembuatan dan penyusunan RPP dan silabus, pembuatan dan penggunaan metode dan media pembelajaran. Pada kategori berhasil KTSP memperoleh persentase lebih besar daripada Kurikulum 2013 dengan persentase 14% untuk KTSP dan 6% untuk Kurikulum 2013. Selanjutnya pada kategori kurang berhasil Kurikulum 2013 memperoleh persentase lebih besar daripada KTSP dengan persentase 6% untuk Kurikulum 2013 dan 1% untuk KTSP. Kekurangberhasilan terdapat pada indikator semangat guru dalam memperkaya bahan ajar. Pada kategori tidak berhasil keduanya memperoleh persentase sebesar 0% atau tidak ada responden yang menyatakan bahwa KTSP dan Kurikulum 2013 tidak berhasil.

Selanjutnya, hasil analisis evaluasi proses berdasarkan kuesioner yang dibagikan kepada siswa dapat dilihat pada Gambar 6.



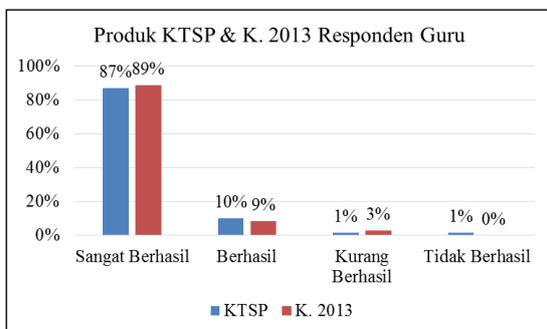
Gambar 6. Penyebaran Jawaban Proses KTSP & Kurikulum 2013 (responden siswa)

Hasil tersebut menunjukkan bahwa implementasi KTSP dan Kurikulum 2013 pada aspek proses termasuk pada kategori sangat berhasil. KTSP memperoleh persentase lebih besar daripada Kurikulum 2013 dengan persentase 75% untuk KTSP dan 71% untuk Kurikulum 2013. Keberhasilan terdapat pada indikator pemanfaatan dan penggunaan media pembelajaran untuk membuat proses belajar lebih baik dan menyenangkan. Pada kategori berhasil KTSP memperoleh persentase lebih kecil daripada Kurikulum 2013 dengan persentase 22% untuk KTSP dan 25% untuk Kurikulum 2013. Pada kategori kurang berhasil Kurikulum 2013 memperoleh persentase lebih besar daripada KTSP dengan persentase 5% untuk Kurikulum 2013 dan 4% untuk KTSP. Kekurangberhasilan ditunjukkan pada indikator siswa lebih inovatif dan produktif. Sedangkan untuk kategori tidak berhasil kedua kurikulum memperoleh persentase sebesar 0% atau tidak ada responden yang menyatakan bahwa KTSP dan Kurikulum 2013 tidak berhasil.

Hasil penelitian ini menunjukkan bahwa implementasi KTSP dan Kurikulum 2013 pada aspek proses termasuk dalam kategori sangat berhasil untuk responden guru dan siswa. Itu berarti indikator komponen proses terpenuhi yakni pembuatan dan penyusunan Rancangan Program Pembelajaran (RPP) dan silabus, penggunaan metode dan media pembelajaran, pemberian motivasi dan pelaksanaan evaluasi pembelajaran.

Evaluasi Produk KTSP dan Kurikulum 2013

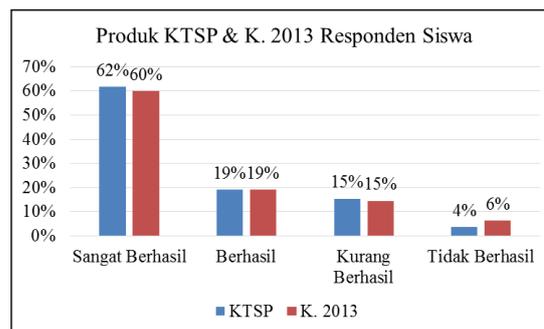
Evaluasi produk diperoleh dari analisis terhadap produk implementasi KTSP dan Kurikulum 2013. Hasil penilaian komponen produk dengan responden guru pada KTSP dan Kurikulum 2013 ditunjukkan melalui Gambar 7.



Gambar 7. Sebaran Jawaban Aspek Produk KTSP & Kurikulum 2013 (responden guru)

Hasil tersebut menunjukkan bahwa implementasi KTSP dan Kurikulum 2013 pada aspek produk termasuk dalam kategori sangat berhasil. KTSP mencapai 87% dan Kurikulum 2013 sebesar 89%. Keberhasilan tertinggi dicapai pada indikator produk silabus dan RPP yang dihasilkan oleh guru, meningkatnya prestasi akademik dan non-akademik. Pada kategori berhasil KTSP memperoleh persentase lebih besar daripada Kurikulum 2013 dengan persentase 10% dan Kurikulum 2013 sebesar 9%. Pada kategori kurang berhasil Kurikulum 2013 memperoleh persentase lebih besar dari. Kekurangberhasilan ini terdapat pada indikator kreativitas guru dalam penggunaan media pembelajaran. Sedangkan untuk kategori tidak berhasil KTSP memperoleh persentase sebesar 1% dan Kurikulum 2013 memperoleh persentase sebesar 0% atau tidak ada responden yang menyatakan bahwa Kurikulum 2013 tidak berhasil.

Selanjutnya, hasil analisis evaluasi produk berdasarkan kuesioner yang dibagikan kepada siswa dapat dilihat pada Gambar 8.



Gambar 8. Sebaran Jawaban Aspek Produk KTSP & Kurikulum 2013 (responden siswa)

Hasil tersebut menunjukkan bahwa implementasi KTSP dan Kurikulum 2013 pada aspek produk termasuk dalam kategori sangat berhasil. KTSP memperoleh persentase lebih besar daripada Kurikulum 2013 dengan persentase 62% dan Kurikulum 2013 sebesar 60%. Keberhasilan ini dicapai pada indikator guru menyajikan pelajaran lebih lengkap dan mendetail. Pada kategori berhasil kedua kurikulum memperoleh persentase yang sama besar yakni 19%. Pada kategori kurang berhasil kedua kurikulum memperoleh persentase yang sama besar yakni 15% yaitu pada indikator mendorong peningkatan sarana pendidikan. Pada untuk kategori tidak berhasil KTSP memperoleh persentase lebih kecil daripada Kurikulum 2013 dengan persentase 4% untuk KTSP dan 6% untuk Kurikulum 2013.

Hasil penelitian ini menunjukkan bahwa implementasi KTSP dan Kurikulum 2013 pada aspek produk termasuk dalam kategori sangat berhasil untuk responden guru dan siswa. Hal tersebut berarti bahwa indikator yang menunjukkan keberhasilan komponen produk terpenuhi. Indikator penunjuk keberhasilan komponen ini adalah karya yang dihasilkan guru berupa RPP, silabus, kalender pendidikan. Selain itu produk lain yang dihasilkan adalah prestasi akademik dan non-akademik siswa serta adanya peningkatan sarana dan prasarana pendidikan pada setiap SMK se-Kabupaten Belu.

Hambatan dalam Implementasi KTSP & Kurikulum 2013

Berdasarkan hasil analisis data, diketahui bahwa faktor penghambat dalam implementasi KTSP pada pembelajaran di SMK se-Kabupaten Belu adalah sebagai berikut. *Pertama*, KTSP menuntut guru untuk melaksanakan sistem penilaian secara mandiri dan berkelanjutan, namun dalam praktiknya guru belum mampu memenuhi tuntutan ini. Guru mengalami hambatan dalam proses penilaian karena perbedaan karakteristik peserta didik sehingga sulit untuk mengidentifikasi atau menghafal setiap peserta didik. *Kedua*, Guru dituntut untuk menggunakan metode pembelajaran yang bervariasi dan menyenangkan seperti: metode *inquiry, discovery, kontekstual, problem solving* dan sebagainya, namun dalam pelaksanaannya beberapa guru mengalami hambatan seperti keterbatasan waktu dan kurangnya penguasaan teknologi. *Ketiga*, mental siswa yang kurang siap dan untuk mandiri dalam belajar.

Hambatan-hambatan tersebut tidak membuat guru-guru SMK di Kabupaten Belu berdiam diri. Berdasarkan hasil analisis kuesioner terbuka yang diberikan kepada guru diperoleh beberapa cara yang dilakukan untuk mengatasi hambatan dalam proses implementasi kurikulum. Cara-cara yang dilakukan guru dalam mengatasi hambatan dari proses implementasi KTSP pada pembelajaran di SMK yaitu (a) mencari sumber belajar lain, (b) mengikuti pelatihan seperti diklat, dan (c) menyediakan waktu tambahan untuk pelajaran yang tertinggal.

Selanjutnya faktor penghambat dalam implementasi Kurikulum 2013 pada pembelajaran di SMK se-Kabupaten Belu diketahui sebagai berikut. *Pertama*, guru belum siap dan sulit mengubah pola pikir. Keberhasilan pembelajaran juga tergantung pada kesiapan guru melalui sosialisasi, pelatihan, dan diklat. Pelatihan meliputi pemilihan instruktur nasional, guru inti, guru kelas dan guru mata pelajaran. Guru-guru yang tidak didampingi dan siapakan dengan baik akan cenderung menggunakan metode lama seperti metode ceramah. *Kedua*, guru pada be-

berapa mata pelajaran kehilangan jam mengajar. Padahal guru terikat dengan syarat mengajar 24 jam tiap minggu. Karena itu meniadakan dan menggabungkan beberapa mata pelajaran menjadi keresahan bagi guru.

Ketiga, minimnya informasi mengenai pedoman dan sosialisasi Kurikulum 2013 dan buku pedoman. Kurangnya sosialisasi kurikulum kepada kepala program keahlian di SMK membingungkan pihak sekolah, guru dan murid. Di samping itu terdapat banyak kasus kekurangan buku panduan pelajaran dari pemerintah pusat pada satuan pendidikan karena belum didistribusikan dengan baik.

Hal yang dilakukan guru untuk mengatasi hambatan dalam proses implementasi Kurikulum 2013 pada pembelajaran di SMK se-Kabupaten Belu antara lain: (a) berusaha membuat format penilaian secara baik dan lebih sederhana, (b) mengatur waktu pembelajaran, (c) melengkapi sarana dan prasarana yang belum tersedia, (d) mengikuti sosialisasi, diklat, dan pelatihan lebih lanjut tentang implementasi Kurikulum 2013, (e) menambah wawasan dengan banyak membaca buku, dan (f) menyesuaikan model pembelajaran dengan materi yang diajarkan.

Kelebihan dan Kekurangan KTSP dan Kurikulum 2013

Kelebihan dan Kekurangan KTSP

Berdasarkan angket terbuka diperoleh data tentang kelebihan dari implementasi KTSP pada pembelajaran di SMK se-Kabupaten Belu. Kelebihan KTSP adalah (a) fleksibel (dapat disesuaikan dengan kondisi sekolah) dan memberikan peluang bagi sekolah untuk mengembangkan sendiri kurikulum yang sesuai kebutuhan, (b) membangun kemandirian dan meningkatkan kreativitas siswa, guru, dan sekolah, (c) guru memaknai sepenuhnya materi yang diajarkan, (d) penyebaran materi merata sehingga mudah dipahami dan diimplementasikan guru, (e) sistem penilaiannya sederhana, dan (f) pembelajaran yang terpusat kepada siswa membuat siswa lebih aktif. Sedangkan kekurangan KTSP dalam pembelajaran di SMK se-Kabupaten Belu adalah, (a) kekurangan

buku dan sumber belajar, (b) alokasi waktu belum sesuai, (c) kurangnya sarana dan prasarana pendukung, (d) kurangnya SDM, (e) kurang maksimal mencapai hasil non-akademik (siswa), dan (f) guru lebih aktif daripada siswa dalam proses pembelajaran.

Kelebihan dan Kekurangan Kurikulum 2013

Berdasarkan hasil analisis penelitian ini diketahui bahwa implementasi Kurikulum 2013 pada pembelajaran di SMK se-Kabupaten Belu memiliki kelebihan dan kekurangan. Kelebihan-kelebihan Kurikulum 2013 adalah: (a) ketersediaan silabus dan buku pengesahan bagi guru, (b) kemudahan dalam menyusun RPP, (c) siswa sebagai pusat pembelajaran (*student center*), dan (d) mendorong peningkatan kreativitas guru dan siswa. Sedangkan kekurangan implementasi Kurikulum 2013 adalah, (a) penilaian yang terlalu rumit, (b) kurangnya sosialisasi dan pelatihan untuk guru, (c) terbatasnya buku dan literatur, (d) kurangnya fasilitas, sarana dan prasarana penunjang dalam proses pembelajaran, dan (e) belum semua guru mampu dan paham untuk mengimplementasikan Kurikulum 2013. Dari kekurangan-kekurangan di atas, yang paling dominan adalah belum semua guru mampu dan paham untuk mengimplementasikan Kurikulum 2013.

Simpulan

Berdasarkan rumusan masalah, uraian teoretis dan analisis data hasil penelitian dapat ditarik beberapa simpulan sebagai berikut. Pertama, implementasi KTSP dan Kurikulum 2013 pada SMK se-Kabupaten Belu untuk aspek konteks masuk dalam kategori sangat berhasil. Kedua, implementasi KTSP dan Kurikulum 2013 pada SMK se-Kabupaten Belu untuk aspek input termasuk kategori sangat berhasil, sedangkan Kurikulum 2013 untuk responden siswa masuk dalam kategori berhasil.

Ketiga, implementasi KTSP dan Kurikulum 2013 pada SMK se-Kabupaten Belu untuk kriteria proses termasuk dalam kategori sangat berhasil. Keempat, Implementasi KTSP dan Kurikulum 2013 pada SMK se-

Kabupaten Belu untuk aspek produk termasuk kategori sangat berhasil dan Kurikulum 2013 masuk dalam kategori berhasil.

Kelima, hambatan paling dominan yang dialami dalam implementasi KTSP adalah guru belum secara maksimal memenuhi tuntutan untuk melaksanakan sistem penilaian secara mandiri dan berkelanjutan, serta belum mahir menggunakan metode pembelajaran yang variatif dan menyenangkan. Hambatan ini diatasi dengan aktif mencari sumber belajar di luar bahan yang disediakan oleh pemerintah. Keenam, hambatan utama dari implementasi Kurikulum 2013 adalah format penilaian yang belum maksimal dipahami oleh guru SMK. Hambatan ini diatasi dengan cara guru mencari format penilaian yang lebih sederhana, berusaha untuk mengikuti setiap sosialisasi dan diklat yang berhubungan dengan Kurikulum 2013.

Ketujuh, kelebihan dari implementasi KTSP yang paling dominan adalah fleksibel dan memberikan peluang bagi sekolah untuk mengembangkan kurikulum yang sesuai kebutuhan sedangkan kekurangannya adalah keterbatasan buku dan sumber literatur belajar. Kedelapan, kelebihan Kurikulum 2013 paling dominan adalah meningkatkan kreativitas guru dan siswa sedangkan kekurangannya adalah belum semua guru memahami dan mampu menggunakan Kurikulum 2013.

Hasil simpulan penelitian ini dapat digunakan bahan masukan kepada Dinas Pendidikan Kabupaten Belu. Pertama, Kepada Dinas Pendidikan, Pemuda dan Olahraga Kabupaten Belu untuk mempertimbangkan menggunakan kembali Kurikulum 2013 dalam proses pembelajaran. Selain itu perlu untuk meningkatkan sosialisasi kurikulum bagi para guru dan membuat pelatihan-pelatihan yang berkaitan dengan kurikulum karena masih ada guru yang mengalami kesulitan dalam membuat penilaian. Selain itu, Dinas Pendidikan perlu untuk meningkatkan sarana dan prasarana pendidikan di beberapa SMK dengan persediaan yang belum memadai. Di samping itu perlu pemerataan penempatan tenaga guru di SMK.

Kedua, Kepada sekolah dan guru supaya meningkatkan pengetahuan dan keterampilan dalam menyusun silabus, RPP, dan menggunakan metode pembelajaran yang variatif serta menyenangkan. Ketiga, hasil penelitian ini menjadi masukan untuk memperbaiki kebijakan-kebijakan yang berhubungan dengan implementasi kurikulum agar pelaksanaannya dapat berjalan baik sesuai pedoman sehingga mampu meningkatkan kualitas proses pembelajaran di SMK di Kabupaten Belu.

Daftar Pustaka

- Depdiknas. Undang-Undang Nomor 20 Tahun 2003 tentang Sistem Pendidikan Nasional (2003). Jakarta.
- Mardapi, D. (2012). *Pengukuran, penilaian dan evaluasi pendidikan*. Yogyakarta: Nuha Litera.
- Menteri Pendidikan dan Kebudayaan Republik Indonesia. Peraturan Menteri Pendidikan dan Kebudayaan Republik Indonesia Nomor 70 Tahun 2013 tentang Kerangka Dasar dan Struktur Kurikulum Sekolah Menengah Kejuruan/Madrasah Aliyah Kejuruan (2013).
- Mulyasa, E. (2006). *Kurikulum tingkat satuan pendidikan, konsep, karakteristik dan Implementasi*. (Rosdakarya, Ed.). Bandung.
- Muslich, M. (2007). *KTSP: pembelajaran berbasis kompetensi dan kontekstual*. Jakarta: PT. Bumi Aksara.
- Stufflebeam, D. L., Madaus, G. F., & Kellaghan, T. (Eds.). (2002). *Evaluation models. Viewpoints on educational and human services evaluation* (Vol. 49). Dordrecht: Kluwer Academic Publishers.
<https://doi.org/10.1007/0-306-47559-6>
- Stufflebeam, D. L., & Shinkfield, A. J. (2014). *Evaluation, teori, models and application*. San Francisco: John Wiley & Sons, Inc.
- Susilo, M. J. (2007). *Kurikulum tingkat satuan pendidikan*. Yogyakarta: Pustaka Pelajar.

ANALISIS KUALITAS SOAL DI PERGURUAN TINGGI BERBASIS APLIKASI *TAP*

Akbar Iskandar^{1*}, Muhammad Rizal¹

¹STMIK AKBA Makassar

¹Jln. Perintis Kemerdekaan, Tamalanrea, Tamalanrea Jaya, Makassar, Sulsel, Indonesia

* Corresponding Author. Email: akbar.iskandar06@gmail.com

Abstrak

Tujuan penelitian ini adalah untuk menemukan butir instrumen yang berkualitas berdasarkan tingkat validitas, reliabilitas, tingkat kesukaran, daya beda dan pengecoh. Jenis penelitian ini adalah jenis penelitian *ex post facto*. Objek penelitian semua jawaban hasil tes calon mahasiswa baru tahun 2014-2016. Data dikumpulkan dengan metode Observasi, wawancara dan dokumentasi. Teknik analisis data secara kualitatif dan kuantitatif. Hasil penelitian menunjukkan bahwa untuk validitas isi ditemukan nilai validitas isi (*vi*) sebesar 0,42 termasuk kategori sedang. Selanjutnya tampak bahwa nilai koefisien reliabilitas instrumen sebesar 0,514. Jumlah butir soal yang sukar sebanyak 57,5%, kategori sedang sebanyak 42,5% dan tidak terdapat soal kategori mudah. Selain itu, butir soal yang memiliki daya beda sangat baik sebanyak 5%, baik sebanyak 20%, perlu revisi sebanyak 13,75%, tidak baik sebanyak 61,25%. Sedangkan option yang tidak berfungsi dengan baik pada saat dijadikan sebagai pengecoh sebanyak 5 butir soal, akan tetapi terdapat 40 butir soal yang harus direvisi karena option pengecoh malah dianggap sebagai kunci jawaban oleh peserta yang pintar.

Kata kunci: *validitas, reliabilitas, tingkat kesukaran, daya beda, pengecoh, TAP*

ANALYSIS OF EXAMINATION INSTRUMENTS QUALITY AT UNIVERSITY BASED ON *TAP* APPLICATION

Abstract

This study is aimed at finding the items of quality instruments based on the level of validity, reliability, difficulty, differentiation, and distraction. This study was an *ex post facto* research. The object of research was all the answers of enrollment test results of prospective students in 2014-2016. The data were collected by using observation, interview, and documentation methods, and the data were analyzed using both qualitative and quantitative technique. The results of the research show that the value of the content validity found is 0.42, which belongs to medium category. Furthermore, it appears that the value of the instrument reliability coefficient is 0.514. The number of the question items which belongs to difficult category is 57.5%, medium category is 42.5%, and no item is in easy category. In addition, the percentage of the question items that belong to the category of very good, good, need a little revision, and not good, in terms of the items' differentiation, are respectively 5%, 20%, 13.75%, and 61.25%. Further, the options that do not work properly when used as distractors are 5 items, but there are 40 items that must be revised because the distractive option is even considered as the key answer by smart participants.

Keywords: *validity, reliability, difficulty level, differentiation, distraction, TAP*

Permalink/DOI: <http://dx.doi.org/10.21831/pep.v22i1.15609>

Pendahuluan

Aplikasi *Test Analysis Program* (TAP) merupakan salah satu program yang dapat digunakan dalam bidang pengukuran untuk menganalisis kualitas butir sebuah instrumen. Hasil analisisnya dapat dijadikan sebagai sumber informasi akurat dan sebagai dasar pengambilan keputusan, apakah instrumen tersebut baik atau tidak.

Salah satu penyebab rusaknya mutu pendidikan adalah hasil tes masuk yang tidak akurat. Untuk itu, penilaian yang benar akan memberikan informasi yang tepat serta mendorong dalam meningkatkan motivasi dan prestasi dalam pembelajaran mahasiswa. Hal ini dijelaskan oleh Iskandar (2013, p. 37) yang menyatakan bahwa sistem tes dan penilaian yang baik akan mendorong mahasiswa dalam meningkatkan motivasi dan prestasi dalam pembelajaran.

Namun yang sering terjadi dalam dunia pendidikan, kita sering dihadapkan pada masalah pengambilan keputusan, apakah seorang mahasiswa harus mengulang materi tertentu, pantas lulus ataukah harus tidak lulus. Hal tersebut bukanlah pekerjaan yang mudah. Dibutuhkan pertimbangan yang matang agar dapat menghasilkan suatu keputusan yang benar dan tepat sehingga tidak merugikan mahasiswa. Untuk itu, keputusan yang tepat dan benar sangat dipengaruhi oleh kualitas instrumen yang digunakan. Jika kualitas instrumen jelek maka pengambilan keputusan juga dipastikan akan jelek.

Instrumen merupakan suatu alat yang digunakan untuk mengukur suatu obyek penelitian, oleh karena itu instrumen tersebut harus memenuhi kriteria yang baik. Persyaratan instrumen yang baik setidaknya memenuhi syarat valid dan reliabel. Disamping memenuhi syarat valid dan reliabel juga harus memperhatikan karakteristik butir yaitu tingkat kesukaran, daya beda, dan keberfungsian pengecoh. Hal tersebut sesuai dengan pendapat Mansyur, Rasyid, & Suratno (2015, p. 30) yang mengatakan bahwa untuk memperoleh informasi yang akurat maka dibutuhkan instrumen yang sah dan handal.

Menurut (Sudrajat, 2008) penilaian (*assessment*) adalah penerapan berbagai cara

dan penggunaan beragam alat penilaian untuk memperoleh informasi tentang sejauh mana hasil belajar peserta didik atau ketercapaian kompetensi (rangkaiannya kemampuan) peserta didik. Sehingga hasil dari proses penilaian melahirkan keputusan-keputusan yang berkaitan dengan mahasiswa meliputi penempatan mahasiswa pada program pendidikan yang berbeda, pemberian nilai pada mahasiswa, membimbing dan mengarahkan mahasiswa, pemilihan mahasiswa untuk mengikuti program-program pendidikan, pemberian penghargaan dan sertifikat terhadap kompetensi mahasiswa.

Kata tes berasal dari bahasa latin *testum*, yang berarti alat untuk mengukur tanah. Sehingga Tes dapat didefinisikan sebagai sejumlah pertanyaan yang membutuhkan jawaban atau sejumlah pernyataan yang harus diberikan tanggapan guna mengukur tingkat kemampuan seseorang atau mengungkap aspek tertentu dari orang yang dikenai tes.

Tes didefinisikan sebagai suatu instrumen atau prosedur sistematis untuk mengobservasi dan menjelaskan satu atau beberapa karakteristik siswa dengan menggunakan suatu skala numerik atau skema klasifikasi (Nitko & Brookhart, 2007, p. 7). Selanjutnya Sax (1980, p. 13) berpendapat bahwa "*a test may be defined as a task or series of tasks used to obtain systematic observations presumed to be representative of educational or psychological traits or attributes*".

Analisis butir tes pada umumnya dimaksudkan untuk mengetahui besar kecilnya indeks tingkat kesulitan, indeks daya beda dan efektivitas pengecoh butir-butir soal yang bersangkutan. Analisis tes dapat dilakukan dengan menggunakan salah satu dari dua cara, tergantung teori tes mana yang digunakan. Teori tes tersebut dapat berupa teori tes klasik atau teori tes modern (Suryabrata, 2002, p. 24).

Analisis kualitas tes merupakan suatu tahap yang harus ditempuh untuk mengetahui derajat kualitas suatu tes, baik secara keseluruhan maupun butir soal yang menjadi bagian dari tes tersebut. Dalam penilaian hasil belajar, tes diharapkan dapat meng-

gambarkan sampel perilaku dan menghasilkan nilai yang objektif serta akurat. Jika tes yang digunakan dosen kurang baik, maka hasil yang diperoleh pun tentunya kurang baik pula. Hal ini dapat merugikan mahasiswa itu sendiri, artinya hasil yang diperoleh mahasiswa menjadi tidak objektif. Oleh sebab itu, tes yang digunakan harus memiliki kualitas yang baik. Tes hendaknya disusun berdasarkan prinsip dan prosedur penyusunan tes. Setelah digunakan perlu diketahui apakah tes tersebut berkualitas baik atau tidak maka perlu dilakukan analisis kualitas tes (Arifin, 2012, p. 22).

Hasil observasi pendahuluan pada lokasi penelitian ditemukan bahwa soal yang digunakan dalam penerimaan calon mahasiswa baru tidak melalui analisis secara empirik (uji validitas, reliabilitas, tingkat kesukaran, daya beda, dan pengecoh). Hal tersebut diungkapkan salah seorang petugas pelaksana tes. Sehingga bisa dipastikan bahwa informasi yang dikumpulkan dari tes yang diberikan, mengandung bias atau tidak sesuai dengan apa yang diinginkan.

Salah satu syarat instrumen tes yang baik harus melihat daya beda (diskriminasi) suatu butir tes dimana daya beda ini digunakan untuk membedakan antara peserta tes yang berkemampuan tinggi dan berkemampuan rendah. Daya beda butir dapat diketahui dengan melihat besar kecilnya indeks diskriminasi. Adapun fungsi dari daya pembeda tersebut adalah mendeteksi perbedaan individual yang sekecil-kecilnya di antara para peserta tes. Ramdani (2012, p. 28) juga mengungkapkan daya pembeda sebuah soal bertujuan untuk menunjukkan kemampuan soal tersebut atau membedakan antara mahasiswa yang pandai dengan yang kurang pandai.

Suatu soal yang dapat dijawab benar oleh mahasiswa pandai maupun oleh mahasiswa kurang pandai, maka soal itu tidak baik karena tidak mempunyai daya beda. Demikian pula jika semua mahasiswa, baik pandai maupun kurang pandai tidak dapat menjawab dengan benar maka soal tersebut juga tidak memiliki daya pembeda. Soal yang baik dan mempunyai daya pembeda adalah

soal yang dapat dijawab benar oleh siswa-siswa yang pandai saja (Arikunto, 2009, p. 26). Butir soal yang tidak memiliki daya pembeda diduga terlalu mudah atau terlalu sulit maka perlu diperbaiki atau diganti dengan pertanyaan lain.

Penentuan daya beda butir biasanya dilakukan dengan menggunakan indeks korelasi, diskriminasi, dan indeks keselarasan item. Dari ketiga cara tersebut yang paling sering digunakan adalah indeks korelasi. Ada dua macam teknik korelasi yang biasa digunakan untuk menghitung nilai daya beda, yaitu: (1) teknik *point biserial*, (2) teknik *biserial* (Mansyur, Rasyid, & Suratno, 2009, p. 155). Untuk memudahkan perhitungan daya pembeda butir soal, Suprananta (2004) (Mansyur & Rasyid, 2007, p. 161) memberikan formula umum dengan rumus sebagai berikut :

$$D = \frac{\Sigma A}{n_A} - \frac{\Sigma B}{n_B}$$

Keterangan :

D : Indeks daya beda butir soal

ΣX_A : Banyaknya peserta tes yang menjawab benar kelompok atas

ΣX_B : Banyaknya peserta tes yang menjawab benar kelompok bawah

n_A : Banyaknya peserta tes pada kelompok atas

n_B : Banyaknya peserta tes pada kelompok bawah

Dari rumus di atas dapat dimaknai bahwa daya beda adalah perbedaan antara proporsi kelompok atas yang menjawab benar butir tes dengan proporsi kelompok bawah yang menjawab benar butir tes. Rumus tersebut dapat digunakan untuk menghitung daya beda butir soal dalam bentuk pilihan ganda.

Koefisien daya beda butir soal bergerak dari -1,00 sampai +1,00 maksudnya adalah jika suatu butir memiliki korelasi negatif, maka dapat dikatakan bahwa butir tersebut menyesatkan, karena subjek yang terdiri dari kelompok pandai menjawab salah soal yang ada daripada subjek pada kelompok kurang

pandai, sehingga harus didrop atau dibuang. Untuk menyatakan bahwa besaran daya beda dapat berfungsi dengan baik, ada beberapa patokan yang dapat digunakan.

Butir soal yang diterima harus memiliki indeks daya beda > 0,30 atau lebih. Butir dengan indeks daya beda kurang dari antara 0,10 sampai 0,30 perlu direvisi, dan jika daya bedanya < 0,10 maka butir tersebut harus dibuang. Sejalan dengan hal ini, Crocker & Algina (1986) (Mansyur & Rasyid, 2007, p. 155) memberikan patokan indeks daya beda seperti Tabel 1.

Tabel 1. Indeks Daya Beda Butir

Indeks Daya Beda	Kriteria Butir
$0,40 \leq D \leq 1,0$	Sangat Baik
$0,3 \leq D < 0,4$	Baik
$0,2 \leq D < 0,3$	Cukup dan perlu sedikit revisi
$D < 0,2$	Tidak Baik

Selain syarat daya beda juga harus memperhatikan tingkat kesukaran butir karena soal yang baik adalah soal yang tidak terlalu mudah atau tidak terlalu sukar. Soal yang terlalu mudah tidak merangsang mahasiswa untuk mempertinggi usaha memecahkannya. Sebaliknya, soal yang terlalu sukar akan menyebabkan mahasiswa berputus asa dan tidak mempunyai semangat untuk mencoba lagi karena berada di luar jangkauannya. Bilangan yang menunjukkan sukar dan mudahnya suatu soal disebut dengan indeks kesukaran. Adapun besarnya indeks kesukaran adalah antara 0,00 sampai dengan 1,00 (Mansyur et al., 2009, p. 20; Rofiah, Aminah, & Ekawati, 2013, p. 4). Indeks kesukaran ini menunjukkan taraf kesukaran soal. Soal dengan indeks kesukaran 0,00 menunjukkan bahwa soal itu terlalu sukar, sebaliknya indeks 1,00 menunjukkan bahwa soalnya terlalu mudah.

Proporsi menjawab benar p (*proportion correct*) adalah indeks kesukaran soal yang paling sederhana dan sering digunakan dalam menentukan besaran indeks. Rumus untuk menentukan besarnya indeks kesukaran secara matematis dirumuskan oleh Mansyur et al. (2009, p. 21) sebagai berikut:

$$P_i = \frac{\sum X_i}{S_{mi} \cdot N}$$

- P_i : tingkat kesukaran butir soal
 $\sum X_i$: Jumlah peserta tes yang menjawab benar
 S_{mi} : Skor maksimum
 N : Jumlah peserta tes

Kriteria yang digunakan untuk menentukan jenis tingkat kesukaran butir soal disajikan dalam Tabel 2.

Tabel 2. Tingkat Kesukaran Butir Soal

Nilai P	Kategori
$p < 0,30$	Sukar
$0,30 \leq p \leq 0,70$	Sedang
$p > 0,70$	Mudah

Dari penjelasan di atas ada beberapa hal yang bisa disimpulkan berkaitan dengan indeks kesukaran butir yaitu bahwa nilai p bagi suatu butir hanya menunjukkan indeks bagi kelompok yang diuji. Harga p ini bisa berubah jika tes diujikan pada kelompok yang berbeda. Selain itu, indeks kesukaran yang dihasilkan dari rumus ini adalah indeks kesukaran yang berlaku bagi kelompok secara keseluruhan, bukan perorangan. Indeks kesukaran bagi tiap peserta tes tidak bisa disimpulkan dengan melihat indeks proporsi menjawab benar p .

Setiap tes pilihan ganda memiliki satu pertanyaan serta beberapa pilihan jawaban. Di antara pilihan jawaban yang ada, hanya satu yang benar. Selain jawaban yang benar tersebut, juga ada jawaban salah, yang dikenal dengan *distractor* (pengecoh). Dengan demikian, efektivitas pengecoh adalah seberapa baik pilihan yang salah tersebut dapat mengecoh peserta tes yang memang tidak mengetahui kunci jawaban yang tersedia. Semakin banyak peserta tes yang memilih pengecoh tersebut, maka distaktor itu dapat menjalankan fungsinya dengan baik. Kriteria pengecoh yang baik adalah apabila pengecoh tersebut dipilih oleh paling sedikit 5% dari peserta tes (Hamzah & Koni, 2012, p. 120).

Distractor (pengecoh) berfungsi untuk mengidentifikasi peserta tes yang berkemampuan tinggi. Pengecoh dikatakan berfungsi efektif apabila dipilih lebih banyak oleh peserta tes yang berasal dari kelompok bawah (berkemampuan rendah), sebaliknya apabila pengecoh itu dipilih lebih banyak oleh peserta tes yang mempunyai kemampuan tinggi, maka pengecoh itu tidak berfungsi sebagaimana mestinya. Bila pengecoh dipilih secara merata, maka termasuk pengecoh yang baik, apabila pengecoh lebih banyak dipilih oleh peserta tes dari kelompok atas dibandingkan dengan kelompok bawah maka termasuk pengecoh yang menyesatkan (Supranata, 2006, p. 32).

Dengan demikian maka pengecoh yang tidak memenuhi kriteria sebagai pengecoh yang baik, karena tidak satupun diantara peserta tes yang memilihnya sebaiknya diganti dengan pengecoh lain yang lebih menarik untuk dipilih oleh peserta tes. Agar semua opsi dalam setiap butir soal dapat berfungsi secara efektif, maka penyusunan pengecoh harus dilakukan sedemikian rupa sehingga tidak terlalu mencolok sebagai opsi yang salah. Pengecoh-pengecoh yang baik adalah yang serupa tetapi tidak sama dengan opsi benar sehingga mempunyai peluang untuk dipilih oleh peserta tes yang tidak berhati-hati.

Metode Penelitian

Penelitian ini termasuk dalam kategori jenis penelitian *Ex Post Facto* yaitu penelitian empiris yang sistematis dimana peneliti tidak mengendalikan variabel bebas secara langsung karena eksistensi dari variabel tersebut telah terjadi, atau karena pada dasarnya variabel tersebut tidak dapat dimanipulasi, Karlinger (Emzir, 2011, p. 18). Penelitian ini dilakukan pada perguruan tinggi STMIK AKBA yang ada di Kota Makassar Propinsi Sulawesi Selatan dengan objek penelitian semua jawaban hasil tes calon mahasiswa baru tahun 2014-2016.

Untuk memperoleh data di lapangan, peneliti menggunakan beberapa teknik yaitu: (1) observasi, yaitu pengumpulan data yang dilakukan dengan cara mengadakan

pengamatan langsung di lapangan, yang ada hubungannya dengan masalah penelitian ini; (2) wawancara, yaitu pengumpulan data yang dilakukan dengan cara mengadakan tanya jawab secara langsung, yang ada hubungannya dengan masalah penelitian ini; (3) dokumentasi, yaitu pengumpulan data melalui referensi-referensi tertulis berupa jawaban hasil tes mahasiswa, buku-buku, bahan ajar, dan lain-lain yang sangat relevan.

Teknik analisis data dilakukan secara kualitatif dan kuantitatif. Analisis kualitatif menggunakan format penelaahan oleh pakar dan analisis secara kuantitatif dengan menampilkan hasil analisis secara klasik yaitu dengan melihat, tingkat kesukaran, daya beda dan efektivitas pengecoh setiap soal atau item melalui program aplikasi TAP.

Adapun kriteria yang digunakan dalam membedakan daya beda merujuk pada pada tabel 1, sedangkan tingkat kesukaran butir soal merujuk pada pada Tabel 2. Pengecoh dinyatakan telah dapat menjalankan fungsinya dengan baik apabila dipilih oleh sekurang-kurangnya 5 % dari seluruh peserta tes (Sudijono, 2009, p. 14). Tingkat validitas tes dapat dilihat dengan mengikuti kriteria validitas isi dari (Gregory, 2007, p. 221) yaitu sebagai berikut.

$0,8 - 1$	= <i>Validitas sangat tinggi</i>
$0,6 - 0,79$	= <i>Validitas tinggi</i>
$0,40 - 0,59$	= <i>Validitas sedang</i>
$0,20 - 0,39$	= <i>Validitas rendah</i>
$0,00 - 0,19$	= <i>Validitas sangat rendah</i>

Selanjutnya kriteria yang digunakan dalam menentukan reliabilitas tes yaitu jika hasil analisis memiliki nilai reliabilitas lebih besar atau sama dengan 0,70 maka dikatakan reliabel (Linn dalam Mansyur et al., 2009, p. 24).

Hasil Penelitian dan Pembahasan

Analisis kualitatif dilakukan untuk mereview butir soal dari aspek materi, konstruksi, dan bahasa sehingga diketahui validitas instrumen tes berdasarkan pandangan para pakar. Aspek-aspek yang diperhatikan dalam memvalidasi instrumen ini adalah:

petunjuk, cakupan soal, bahasa. Tabel 3 adalah rangkuman hasil validasi instrumen tes untuk setiap aspek pengamatan. Dari hasil penilaian para ahli yang berjumlah 2 orang dosen sebagai pakar bidang tes, sehingga instrumen tersebut dapat digunakan dengan melakukan revisi terlebih dahulu.

Tabel 3. Hasil Penilaian Ahli Tes

Rater 1	Rater 2	Hasil Tabulasi Silang
1	3	C
2	2	A
3	3	D
4	2	B
3	3	D
3	2	B
2	3	C

Berdasarkan hasil tabulasi silang 2x2 tersebut maka selanjutnya dimasukkan ke dalam rumus Gregory $V_i = \frac{D}{A+B+C+D}$ maka hasilnya sebesar 0,28. Sehingga instrumen ini memenuhi kriteria validitas isi pada kategori rendah, (Gregory, 2007, p. 221). Hasil penilaian dari 2 orang dosen sebagai pakar IT dapat dilihat seperti pada Tabel 4.

Tabel 4. Hasil Penilaian Ahli IT

Rater 1	Rater 2	Dari Tabulasi Silang
2	1	A
3	2	B
3	4	D
2	3	C
3	3	D
3	2	B
3	3	D

Berdasarkan hasil analisis validitas isi dengan menggunakan rumus Gregory ditemukan nilai V_i sebesar 0,42. Sehingga instrumen ini memenuhi validitas isi kategori sedang. Setelah melewati tahap uji validitas dilanjutkan dengan uji reliabilitas yang bertujuan untuk melihat tingkat kesepakatan validator ahli melalui analisis ICC (*Intraclass Correlation Coefficients*) dan hasil analisis ditemukan sebesar $K=0,514$.

Berdasar pada hasil analisis tersebut tampak bahwa nilai koefisien reliabilitas instrumen ini lebih kecil dari batas bawah reliabilitas yang telah ditentukan yaitu sebesar 0,70 menurut Linn (Mansyur et al., 2009, p. 24), sehingga instrumen tersebut tidak memenuhi kriteria reliabel dan berada pada level reliabilitas cukup (*fair*) ($0,40 \leq K \leq 0,60$: cukup), (Fleiss dalam Widhiarso, 2012, p. 15).

Selanjutnya analisis kuantitatif dilakukan dengan menggunakan Program TAP, yang secara otomatis menganalisis butir instrumen tes seperti tingkat kesukaran, daya beda, efektivitas pengecoh, reliabilitas tes serta beberapa statistik data lainnya (ukuran dari data hasil tes). Hasil analisis secara deskriptif untuk semua butir soal dapat dilihat pada Tabel 5.

Tabel 5. Hasil Analisis Deskriptif Skor Peserta Tes

Kriteria	Hasil analisis
Jumlah peserta tes	300
Kemungkinan skor total	80
Skor maksimal	35
Skor minimum	11
<i>Median</i>	22
<i>Mean</i>	22.33
<i>Standar deviasi</i>	4.546
<i>Variance</i>	20.66

Berdasarkan hasil analisis pada Tabel 5, tampak bahwa jumlah responden yang mengikuti tes ini sebanyak 300 orang. Jika seorang peserta tes menjawab semua soal dengan benar maka skor maksimal yang mungkin diperoleh sebesar 80. Akan tetapi dari hasil tes tersebut, skor maksimal yang diperoleh responden sebesar 35, skor minimal sebesar 11, median 22, mean 22.33, standar deviasi 4.54 dan variance sebesar 20.66. Selanjutnya hasil analisis butir soal secara deskriptif dapat dilihat pada Tabel 6.

Terkait dengan Tabel 6. tampak bahwa jumlah butir soal yang dianalisis sebanyak 80 butir dan tidak ada butir yang hilang. Kemudian dari hasil analisis tersebut ditemukan rerata tingkat kesulitan butir soal

sebesar 0,279 yang menandakan bahwa soal yang digunakan dalam penerimaan calon mahasiswa baru, termasuk dalam kategori sulit, selanjutnya rerata daya beda butir sebesar 0,130. Daya beda dihitung berdasarkan pembagian dua kelompok peserta tes yaitu kelompok atas dan kelompok bawah.

Tabel 6. Hasil Analisis Deskriptif Butir Soal

Kriteria	Hasil Analisis
Jumlah butir soal	80
Rerata tingkat kesulitan butir	0,279
Rerata daya beda butir	0,130
Koefisien reliabilitas (KR20)	0,281
Kesalahan pengukuran	3,83

Kelompok atas dikategorikan sebagai *testee* yang tergolong sebagai anak yang pandai sedangkan peserta tes yang berada pada kelompok bawah dikategorikan sebagai *testee* yang tergolong kurang pandai, sedangkan angka 0,130 tersebut menunjukkan bahwa rata-rata butir soal tidak mampu membedakan antara calon mahasiswa pandai dengan yang kurang pandai. Karena besar kecilnya daya beda dapat diketahui melalui hasil analisis diskriminasi butir dengan membandingkan hasil analisis dengan kriteria yang telah ada.

Lebih lanjut, dengan menggunakan KR 20 diperoleh nilai kesalahan baku pengukuran sebesar 3,83 dan tingkat reliabilitas instrumen tes sebesar 0,281 hal ini menandakan bahwa tingkat keajekan instrumen tes yang digunakan dalam kategori buruk kare-

na indeks reliabilitas lebih kecil dari 0,4 ($K < 0,4$) atau masuk kategori *Bad*, Fleiss (Fleiss dalam Widhiarso, 2012, p. 15). Untuk melengkapi pernyataan sebelumnya, maka hasil analisis tingkat kesukaran setiap butir soal dapat dilihat pada Tabel 7.

Tingkat kesukaran berdasar pada besarnya indeks korelasi yang berkisar antara 0 sampai 1. Makin tinggi indeks korelasi maka butir soal tersebut semakin mudah dan semakin kecil indeks korelasi maka butir soal tersebut semakin sulit. Mansyur et al. (2009, p. 20) membedakan tingkat kesukaran soal ke dalam tiga kategori yaitu soal yang memiliki $p \leq 0,3$ biasanya disebut sebagai soal sukar, soal yang memiliki $p \geq 0,7$ biasanya disebut soal mudah, adapun soal yang memiliki p antara 0,3 sampai 0,7 disebut sebagai soal yang sedang.

Merujuk pada Tabel 7, jumlah butir soal yang masuk kategori sukar sebanyak 46 butir soal (57,5%), butir soal yang berada pada kategori sedang sebanyak 34 butir atau sebesar (42,5%) sedangkan butir soal yang berada pada kategori mudah tidak ada. Dari hasil analisis tersebut kelihatan jumlah soal yang sukar lebih banyak dibandingkan dengan butir soal yang sedang, tetapi perlu diketahui bahwa hasil analisis tersebut bukan satu-satunya indikator bahwa soal yang sukar atau mudah adalah jelek, karena dalam analisis butir soal terdapat beberapa kategori yang harus diperhatikan seperti tingkat kesukaran dan daya beda untuk menilai apakah butir tersebut baik atau tidak. Hasil analisis daya beda dapat dilihat pada Tabel 8.

Tabel 7. Hasil Analisis Tingkat Kesukaran

Kategori	Butir soal	Jumlah
Sukar $P < 0,30$	3, 4, 6, 7, 9, 10, 11, 12, 13, 14, 16, 19, 20, 22, 23, 24, 26, 28, 29, 31, 34, 27, 38, 29, 43, 45, 47, 48, 49, 50, 52, 53, 54, 56, 57, 60, 61, 64, 67, 69, 70, 71, 72, 75, 77, 78.	46 (57,5%)
Sedang $0,30 \leq p \leq 0,70$	1, 2, 5, 8, 15, 17, 18, 21, 25, 27, 30, 32, 33, 35, 36, 40, 41, 42, 44, 46, 51, 55, 58, 59, 62, 63, 65, 66, 68, 73, 74, 76, 79, 80.	34 (42,5%)
Mudah $P > 0,70$	-	-
Total soal	80	80 (100%)

Tabel 8. Hasil Analisis Daya Beda

Kategori	Butir Soal	Jumlah
Sangat baik $0,40 \leq D \leq 1,0$	2, 32, 33, 50	4 (5%)
Baik $0,3 \leq D < 0,4$	8, 9, 24, 25, 30, 35, 36, 45, 48, 52, 54, 63, 68, 74, 79, 80	16 (20%)
Perlu sedikit revisi $0,2 \leq D < 0,3$	1, 15, 17, 42, 44, 51, 55, 56, 58, 62, 66.	11 (13,75%)
Tidak baik $D < 0,2$	3, 4, 5, 6, 7, 10, 11, 12, 13, 14, 16, 18, 19, 20, 21, 22, 23, 26, 27, 28, 29, 31, 34, 37, 38, 39, 40, 41, 43, 46, 47, 49, 53, 57, 59, 60, 61, 64, 65, 67, 69, 70, 71, 72, 73, 75, 76, 77, 78.	49 (61, 25%)
Total soal	80	80 (100%)

Berdasarkan hasil analisis pada Tabel 8, tampak bahwa jumlah butir soal yang memiliki daya beda yang sangat baik sebanyak 4 butir soal (5%) yang berarti bahwa butir soal tersebut mampu membedakan antara peserta tes yang pandai dan peserta tes yang kurang pandai. Selanjutnya, terdapat 20% yang memiliki daya beda pada kategori baik yang berarti butir-butir soal tersebut dapat membedakan antara peserta tes yang pandai dengan peserta tes yang kurang pandai.

Selain butir soal yang memiliki daya beda kategori sangat baik dan kategori baik juga terdapat butir soal yang memiliki daya beda pada kategori cukup, sebanyak 11 butir (13,75%) yang berarti harus melewati tahap revisi. Jika butir-butir soal ini telah direvisi maka butir soal tersebut dapat digunakan. Sedangkan jumlah butir soal yang berada pada kategori tidak baik sebanyak 49 butir yaitu sebesar (61,25%) sehingga harus dibuang.

Selanjutnya, untuk melihat pengecoh butir soal yang berfungsi pada Tabel 9.

Rendahnya daya beda biasanya disebabkan oleh tingkat keberfungsian pengecoh butir soal, selain itu pengecoh juga memberikan dampak terhadap tingkat kesukaran butir soal karena jika terdapat satu atau dua pengecoh pada suatu butir soal yang tidak berfungsi maka indeks tingkat kesukaran

butir soal akan menurun, karena peluang peserta tes untuk menjawab dengan benar semakin meningkat.

Berdasarkan hasil analisis yang tertera pada Tabel 9, tampak bahwa jumlah butir soal yang memiliki pengecoh yang baik untuk option A yaitu sebanyak 68 butir soal dan terdapat 2 butir soal yang memiliki option A tidak berfungsi sebagai pengecoh yaitu butir soal 27, 28 dan 10 butir soal menempatkan option A sebagai kunci. Selanjutnya pada option B terdapat 55 butir soal yang berfungsi sebagai pengecoh dan 25 butir soal yang menempatkan kunci berada pada option B.

Lebih lanjut, terdapat 62 butir soal yang menggunakan option C berfungsi sebagai pengecoh dan 3 butir soal yang menggunakan option C tidak berfungsi dengan baik yaitu pada butir 1, 6, 31 dan 15 butir soal yang menjadikan option C sebagai kunci. Lebih lanjut, terdapat 51 butir soal yang menggunakan option D sebagai pengecoh dan semuanya berfungsi dengan baik. Selain itu, 29 butir soal yang menjadikan option D sebagai kunci. Untuk melengkapi penyajian keberfungsian pengecoh pada Tabel 9 dapat pula dilihat Tabel 10 untuk melihat letak kunci jawaban pada setiap butir soal.

Tabel 9. Keberfungsian Pengecoh

Menjadi kunci	Pengecoh tidak berfungsi dengan baik	Butir soal	Jumlah	Pengecoh yang berfungsi dengan baik	Butir soal	Jumlah
10	A	27, 28.	2	A	1, 2, 3, 5, 6, 8, 9, 11, 12, 13, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 29, 30, 31, 32, 33, 35, 36, 38, 39, 41, 42, 44, 45, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80.	68
25	B	-	-	B	2, 3, 4, 6, 7, 9, 10, 12, 13, 14, 16, 18, 19, 20, 22, 23, 25, 26, 28, 30, 31, 32, 34, 36, 37, 38, 40, 41, 43, 45, 46, 48, 49, 50, 52, 53, 54, 56, 57, 58, 60, 61, 63, 64, 66, 67, 68, 69, 71, 72, 74, 75, 77, 78, 79.	55
15	C	1, 6, 31.	3	C	2, 4, 5, 7, 8, 9, 10, 11, 12, 14, 15, 17, 18, 20, 21, 23, 24, 25, 26, 27, 28, 29, 30, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 50, 51, 52, 54, 55, 56, 58, 59, 60, 62, 63, 65, 66, 68, 70, 71, 73, 74, 76, 77, 79, 80.	62
29	D	-	-	D	1, 3, 4, 5, 7, 8, 10, 11, 13, 14, 15, 16, 17, 19, 21, 22, 24, 26, 27, 29, 31, 33, 34, 35, 37, 39, 40, 42, 43, 44, 46, 47, 49, 51, 53, 55, 57, 59, 61, 62, 64, 65, 67, 69, 70, 72, 73, 75, 76, 78, 80.	51

Tabel 10. Tampilan Kunci Jawaban dalam Aplikasi TAP

```

=====
CORRECT ANSWERS (Item#-Key):
=====
# 1-2 # 2-4 # 3-3 # 4-1 # 5-2 # 6-4 # 7-1 # 8-2 # 9-4 #10-1
#11-2 #12-4 #13-3 #14-1 #15-2 #16-3 #17-2 #18-4 #19-3 #20-4
#21-2 #22-3 #23-4 #24-2 #25-4 #26-1 #27-2 #28-4 #29-2 #30-4
#31-2 #32-4 #33-2 #34-1 #35-2 #36-4 #37-1 #38-4 #39-2 #40-1
#41-4 #42-2 #43-1 #44-2 #45-4 #46-1 #47-2 #48-4 #49-3 #50-4
#51-2 #52-4 #53-3 #54-4 #55-2 #56-4 #57-3 #58-4 #59-2 #60-4
#61-3 #62-2 #63-4 #64-3 #65-2 #66-4 #67-3 #68-4 #69-3 #70-2
#71-4 #72-3 #73-2 #74-4 #75-3 #76-2 #77-4 #78-3 #79-4 #80-2
    
```

Berdasarkan Tabel 10 dalam penyajian kunci jawaban setiap butir soal yang ada pada *CORRECT ANSWERS (Item#-Key)*, seperti pada butir soal 1 letak kunci jawab berada pada option B (2), sedangkan pada soal butir nomor 2 letak kunci jawaban berada pada option D (4) dan seterusnya. Akan tetapi pada setiap peletakan kunci

jawaban pada setiap soal, terkadang ada hal-hal yang menyebabkan letak kunci jawaban tersebut tidak sesuai dengan teori pengembangan tes seperti pada contoh butir soal nomor 3 yang disajikan pada Tabel 11.

Dari hasil analisis pada soal nomor 3 di atas tampak bahwa letak kunci jawaban yang telah ditentukan berada pada option 3

(C), akan tetapi peletakan kunci jawaban yang terletak pada option C tersebut tidak sesuai dan dianggap bahwa option 2 (B) lebih pantas menjadi kunci daripada option 3 (C) karena sebagian besar peserta tes yang pintar menganggap option 2 (B) sebagai kunci jawabannya, berbeda dengan option 3 (C) yang banyak dipilih oleh peserta tes kurang pintar sehingga harus direvisi/dibuang. Selain butir soal nomor 3 juga terdapat butir soal yang lain memiliki hal yang sama, seperti yang tampak pada Tabel 11.

Tampak pada Tabel 11. Jumlah butir soal yang memiliki kunci jawaban yang tidak sesuai yaitu sebanyak 40 butir soal karena pengecoh yang sedianya dijadikan sebagai alat untuk mengelabui peserta tes malah dianggap bisa menjadi kunci jawaban karena pengecoh lebih banyak dipilih oleh peserta tes yang pintar daripada peserta tes yang kurang pintar, sedangkan untuk kunci jawaban dalam setiap butir soal malah banyak dipilih oleh peserta yang kurang pintar.

Tabel 11. Letak Kunci Jawaban

Item	Group	Option 1	Option 2	Option 3	Option 4
3	TOTAL	95 (0.317)	112 (0.373)	46*(0.153)	47 (0.157)
	High	28 (0.337)	41 (0.494)	7 (0.084)	7 (0.084)
	Low	36 (0.343)	28 (0.267)	17 (0.162)	24 (0.229)
	Diff	-8(-0.006)	13#(0.227)	-10(-0.078)	-17(-0.144)

Tabel 12. Pemilihan Option sebagai Letak Kunci Jawaban

Butir soal	Letak kunci (*)	Saran pengganti (#)	Butir soal	Letak kunci (*)	Saran pengganti (#)
4	Option 1	Option 4	46	Option 1	Option 2, 4
6	Option 4	Option 2	47	Option 2	Option 1
7	Option 1	Option 3, 4	49	Option 3	Option 2
10	Option 1	Option 2	53	Option 3	Option 1, 2
11	Option 2	Option 3, 4	57	Option 3	Option 1, 2
13	Option 3	Option 1,2	59	Option 2	Option 4
14	Option 1	Option 4	60	Option 4	Option 2, 3
16	Option 3	Option 4	61	Option 3	Option 4
19	Option 3	Option 2	64	Option 3	Option 1, 2
21	Option 2	Option 1	65	Option 2	Option 4
22	Option 3	Option 1	67	Option 3	Option 1, 2
26	Option 1	Option 2	69	Option 3	Option 2
27	Option 2	Option 1, 4	70	Option 2	Option 4
28	Option 4	Option 2	71	Option 4	Option 2
29	Option 2	Option 4	72	Option 3	Option 4
31	Option 2	Option 1	73	Option 2	Option 1, 3
34	Option 1	Option 4	75	Option 3	Option 2
37	Option 1	Option 4	76	Option 2	Option 3, 4
39	Option 2	Option 1	78	Option 3	Option 2, 4
43	Option 1	Option 3, 4	3	Option 3	Option 2

Simpulan

Berdasarkan hasil analisis dan pembahasan sebelumnya maka dapat disimpulkan bahwa (1) hasil validitas isi ditemukan nilai V_i sebesar 0,42 kategori sedang. (2) indeks reliabilitas sebagai kesepakatan validator ahli melalui analisis ICC diperoleh nilai K sebesar $0,514 < 0,70$, yang berarti berada pada kategori cukup (fair). (3) jumlah butir soal yang sukar sebanyak 57,5%, kategori sedang sebanyak 42,5% dan soal mudah tidak ada.

(4) butir soal yang memiliki daya beda kategori sangat baik sebanyak 5%, kategori baik sebanyak 20%, perlu sedikit revisi sebanyak 13,75%, dan kategori tidak baik sebanyak 61,25% dan (5) pengecoh yang tidak berfungsi dengan baik terdapat pada 5 butir soal, serta terdapat 40 butir soal yang harus direvisi/dibuang karena memiliki pengecoh yang justru dianggap sebagai kunci jawaban oleh peserta yang pintar.

Berdasarkan hasil penelitian dan simpulan yang disampaikan, penelitian memberikan saran sebagai berikut: (1) instrumen tes yang baik sebelum digunakan seharusnya didiskusikan dalam *Focus Group Discussion* (FGD) serta sosialisasi dengan pelaku bidang pendidikan dan ahli dalam bidang ilmu pengukuran dan evaluasi; (2) pemahaman dan kemampuan dalam penyusunan instrumen tes sangat diperlukan untuk menghasilkan butir instrumen yang baik; (3) pelaksana tes harus bersifat adil pada setiap peserta tes.

Daftar Pustaka

- Arifin, Z. (2012). *Evaluasi pembelajaran prinsip, teknik prosedur*. Bandung: PT: Remaja Rosdakarya.
- Arikunto, S. (2009). *Dasar-dasar evaluasi pendidikan*. Jakarta: PT. Bumi Aksara.
- Emzir. (2011). *Metodologi penelitian pendidikan kuantitatif dan kualitatif*. Jakarta: Rajawali Pers.
- Gregory, R. . (2007). *Psychological testing: history, principles, and applications* (5th ed.). New York: Pearson Education Group, Inc.
- Hamzah, B. U., & Koni, S. (2012). *Assesment pembelajaran*. Jakarta: Bumi Aksara.
- Iskandar, A. (2013). Pengembangan perangkat penilaian psikomotor di sekolah menengah kejuruan (SMK). *Inspiration Jurnal Teknologi Informasi Dan Komunikasi*, 3(1). Retrieved from <http://jurnal.akba.ac.id/index.php/inspiration/article/view/30>
- Mansyur, & Rasyid, H. (2007). *Penilaian hasil belajar*. Bandung: Wacana.
- Mansyur, Rasyid, H., & Suratno. (2009). *Assesmen pembelajaran di sekolah*. Yogyakarta: Multi Pressindo.
- Mansyur, Rasyid, H., & Suratno. (2015). *Asesmen pembelajaran disekolah. Panduan bagi guru dan calon guru*. Yogyakarta: Pustaka Pelajar.
- Nitko, A. J., & Brookhart, S. M. (2007). *Educational assessment of students*. New Jersey: Pearson Education.
- Ramdani, Y. (2012). Pengembangan instrumen dan bahan ajar untuk meningkatkan kemampuan komunikasi, penalaran, dan koneksi matematis dalam konsep integral. *Jurnal Penelitian Pendidikan*, 13(1). Retrieved from http://jurnal.upi.edu/file/6-yani_ramdhani.pdf
- Rofiah, E., Aminah, N. S., & Ekawati, E. Y. (2013). Penyusunan instrumen tes kemampuan berpikir tingkat tinggi fisika pada siswa SMP. *Jurnal Pendidikan Fisika*, 1(2).
- Sax, G. (1980). *Principles of educational and psychological measurement and evaluation* (2nd ed.). Belmont: Wadsworth Publishing Company.
- Sudijono, A. (2009). *Pengantar statistik pendidikan*. Jakarta: Rajawali Pers.
- Sudrajat, A. (2008). Pengembangan perangkat penilaian psikomotor. Retrieved January 20, 2012, from <http://akhmadsudrajat.files.wordpress.com>

- com/2008/08/penilaian-
psikomotor.pdf
- Supranata, S. (2006). *Analisis, validitas, reliabilitas dan interpretasi hasil tes*. Bandung: PT. Remaja Rosdakarya.
- Suryabrata, S. (2002). *Pengembangan alat ukur psikologis*. Yogyakarta: Andi.
- Widhiarso, W. (2012). Mengestimasi reliabilitas. Retrieved February 12, 2012, from http://widhiarso.staff.ugm.ac.id/files/bab_2_estimasi_reliabilitas_via_spss.pdf

ESTIMASI KESALAHAN PENGUKURAN *STANDARD SETTING* DALAM PENILAIAN KOMPETENSI MATEMATIKA TINGKAT SMP DI KABUPATEN SUMBAWA

Weni Wendari^{1*}, Samsul Hadi¹

¹Prodi Penelitian & Evaluasi Pendidikan Program Pascasarjana Universitas Negeri Yogyakarta
Jl. Colombo No. 1, Depok, Sleman 55281, Yogyakarta, Indonesia

* Corresponding Author. Email: wendari.weni@gmail.com

Abstrak

Penelitian ini bertujuan untuk mengetahui metode yang lebih akurat dalam mengestimasi kesalahan pengukuran *standard setting* pada metode *Ebel*, *Bookmark*, dan *Contrasting group*. Data penelitian ini merupakan dokumen Dinas Pendidikan dan Kebudayaan Kabupaten Sumbawa berupa respon peserta Ujian Nasional Matematika Paket P0C5520 tahun ajaran 2015/2016 yang berjumlah 352 siswa. Gsuru juga dilibatkan dalam penelitian sebagai panelis dalam *Focus Group Discussion* (FGD). Data yang terkumpul kemudian dianalisis melalui tiga tahap.. Tahap pertama yaitu persiapan, kegiatan pada tahap ini mencakup penyiapan data, penggolongan SMP, dan penentuan karakteristik butir. Tahap kedua yaitu FGD dilakukan dalam dua putaran. Tahap ketiga yaitu mengestimasi kesalahan pengukuran dengan menggunakan pendekatan *Generalizability Theory* dengan bantuan program *eduG*. Hasil penelitian menunjukkan bahwa metode *Contrasting group* memiliki estimasi kesalahan pengukuran paling kecil dibandingkan metode *Ebel* dan *Bookmark*, oleh karena itu, metode *Contrasting group* lebih akurat dibandingkan dengan dua metode lainnya.

Kata kunci: *cut score, standard setting, generalizability theory*

MEASUREMENT ERROR ESTIMATION OF STANDARD SETTING IN MATHEMATICS COMPETENCY ASSESSMENT FOR JUNIOR HIGH SCHOOL IN SUMBAWA REGENCY

Abstract

This research aims to find the most accurate methods in estimating measurement error of standard setting among *Ebel*, *Bookmark*, and *Contrasting group* methods. The data used in this study were 352 students' responses on Mathematics National Exam Package P0C5520 in the academic year of 2015/2016. The document was collected from the Department of Education and Culture in Sumbawa Regency. Teachers were also involved in this research as panelists in the Focus Group Discussion (FGD). The data collected were then analyzed through three stages. The first stage was preparation stage, including the activities of data preparation, school classification, and item characteristics analysis. The second stage was two-round FGD. The third stage was estimating the measurement error using *Generalizability Theory* approach assisted by *eduG* program. The research result shows that *Contrasting Group* method produces the smallest measurement error estimation compared to *Ebel* and *Bookmark* methods, therefore, *Contrasting group* method is considered as the most accurate method.

Keywords: *cut score, standard setting, generalizability theory*

Permalink/DOI: <http://dx.doi.org/10.21831/pep.v22i1.16492>

Pendahuluan

Penilaian dalam kurikulum didesain dengan menggunakan kriteria tertentu. Asumsi yang digunakan pada kriteria yaitu bahwa setiap peserta didik dapat belajar pelajaran apa saja, namun dengan membutuhkan waktu yang berbeda. Kriteria itu berlaku untuk semua peserta didik tanpa membedakan mata pelajaran. Hasil penilaian sering dipandang sebagai tolak ukur keberhasilan peserta didik dalam proses pembelajaran. Hasil penilaian berdasarkan kriteria dikategorikan menjadi dua yaitu lulus dan tidak lulus. Peserta didik dikategorikan lulus jika memenuhi kriteria yang telah ditentukan. Sebaliknya, peserta didik dinyatakan tidak lulus apabila tidak memenuhi kriteria yang telah ditentukan.

Penentuan kriteria kelulusan dapat dilakukan melalui *judgment*. Namun penetapan dengan cara ini memiliki kelemahan, yakni tidak didasarkan pada data empirik dan prosedur yang telah teruji di lapangan. Penetapan dengan *judgment* ini hanya dilakukan berdasarkan pertimbangan dan pendapat yang tidak didasarkan pada data empirik. Kemudian pendapat dan pertimbangan tersebut dijadikan sebagai kebijakan. Sehingga kriteria kelulusan yang ditetapkan dengan cara ini tidak dapat digunakan karena tidak merepresentasikan keadaan yang sebenarnya di lapangan.

Saat ini, Indonesia menggunakan Kurikulum 2013 yang mana kriteria kelulusan untuk Ujian Nasional tidak ditentukan oleh pemerintah. Permendikbud Nomor 5 Tahun 2015 tentang Kriteria Kelulusan Peserta Didik dalam Ujian Nasional (UN) (Menteri Pendidikan dan Kebudayaan Republik Indonesia, 2015) menyatakan bahwa kelulusan peserta didik ditetapkan oleh satuan pendidikan, dalam hal ini kriteria kelulusan ditentukan oleh masing-masing sekolah atau yang biasa dikenal dengan otonomi sekolah. Berarti bahwa setiap sekolah di Indonesia mempunyai kriteria kelulusan yang berbeda-beda, misalnya di Kabupaten Sumbawa terdapat 101 sekolah menengah pertama (SMP) yang terdiri atas 95 SMP Negeri dan 6 SMP Swasta. Masing-masing SMP di

Kabupaten Sumbawa memiliki kriteria kelulusan yang berbeda. Dimana kriteria kelulusan ditentukan hanya berdasarkan *intake* siswa, daya dukung, dan kompleksitas. Ketiga subskor tersebut selanjutnya diambil rata-rata yang kemudian digunakan sebagai batas kelulusan. Penggunaan teknik ini tentu menimbulkan masalah yang cukup serius terutama berkenaan dengan masalah reliabilitas atau keakuratan. Penentuan ketiga komponen tersebut memberikan konsekuensi akan tingginya variabilitas nilai yang mungkin muncul dari para penilai. Permasalahan reliabilitas ini dapat dieliminasi jika batas kelulusan atau kriteria kelulusan ditentukan dengan menggunakan *standard setting*.

Standard setting merupakan proses penentuan suatu titik atau batas dalam skala skor tes tertentu yang digunakan untuk menentukan level performa suatu kebijakan atau untuk membuat klasifikasi (Cizek, 1996, p. 20). *Standard setting* dalam dunia pendidikan banyak digunakan untuk menentukan skor batas kelulusan atau biasa disebut sebagai skor minimum kelulusan. Namun lebih dari itu, *standard setting* juga dapat digunakan sebagai alat bantu untuk memetakan mutu pendidikan, sebab dengan *standard setting* dapat dibuat suatu klasifikasi kompetensi seseorang atau prestasi suatu wilayah tertentu yang termasuk kategori tinggi, sedang, dan rendah. Tidak hanya digunakan dalam menentukan skor minimum kelulusan, *standard setting* juga dapat dimanfaatkan dalam memetakan mutu pendidikan dengan mengklasifikasikan kompetensi peserta didik atau prestasi suatu wilayah tertentu.

Secara garis besar, *standard setting* dibedakan menjadi dua golongan. Pertama yang menggunakan acuan norma dan kedua menggunakan acuan kriteria. Interpretasi nilai beracuan norma yaitu kemampuan peserta tes dibandingkan dengan kemampuan orang lain dalam kelompok acuan. Sementara interpretasi nilai beracuan kriteria, kemampuan peserta tes dibandingkan dengan level kemampuan tertentu.

Sampai saat ini, terdapat sekitar 38 metode yang digunakan dalam menentukan *standard setting* (Berk, 1986, p. 151). Metode

standard setting dibagi kedalam lima kelompok (Zieky, Perie, & Livingston, 2008, p. 86) yaitu (1) metode berdasarkan pertanyaan tes terdiri dari metode *Angoff*, metode estimasi rata-rata, metode *Yes or No Angoff*, metode *Nedelsky*, metode *Ebel*, metode *Bookmark*, dan metode *Item Descriptor Matching*; (2) metode berdasarkan profil skor terdiri dari metode profil performan, dan metode profil *dominant*; (3) metode berdasarkan pada pertimbangan orang atau produk terdiri dari metode *borderline group*, metode *contrasting group*, metode *contrasting group* dengan modifikasi *up and down*, metode *body of work*, dan metode *analytic judgment*; (4) metode berdasarkan pertimbangan kelompok peserta tes yang terdiri dari *Judgments about a Reference Group* dan *Judgments about Two Reference Groups*; dan (5) metode berdasarkan pada kompromi antara pertimbangan absolut dan normatif terdiri dari metode *Beuk* dan metode *Hofstee*.

Retnawati (2014, pp. 165–166) menggolongkan metode *standard setting* menjadi dua kelompok yaitu (1) metode berpusat pada butir/tes, metode ini menggunakan pendekatan klasik dan teori respon butir. Metode berpusat pada butir/tes yang cenderung menggunakan pendekatan klasik terdiri dari metode *Nedelsky*, penilaian profesional, metode *Angoff*, dan metode *Ebel*. Sedangkan metode berpusat pada butir yang menggunakan pendekatan teori respon butir terdiri dari metode *Bookmark* dan metode pemetaan butir (*item mapping*); dan (2) metode yang berpusat pada siswa terdiri dari metode *Contrasting group* dan metode *Borderline*. Penggolongan metode *standard setting* didasarkan pada sudut pandang masing-masing ahli.

Penggolongan berbagai metode *standard setting* dapat mempermudah pemilihan metode *standard setting* yang sesuai dengan karakteristik, tujuan, dan situasi yang terjadi. Pemilihan metode *standard setting* yang tepat akan memberikan kesalahan (*error*) yang kecil. Sehingga penentuan *cut score* akan semakin tepat. Untuk itu sangat penting diperhatikan pemilihan metode *standard setting* yang akan digunakan. Metode *standard setting* yang digunakan akan sangat menentukan

dalam menghasilkan *cut score* yang tepat dengan nilai *error* terkecil.

Cut score merupakan point penting dalam *standard setting*. Penentuan *cut score* bukanlah hal yang mudah (Nudell, 2008, p. 5). *Cut score* yang ditetapkan harus dapat mencerminkan ketercapaian kompetensi minimal yang harus dicapai peserta tes. Ketepatan penentuan *cut score* dalam *standard setting* ditentukan berdasarkan besar kecilnya *error*. Semakin besar nilai *error* maka semakin tidak tepat penentuan *cut score*. Sebaliknya, semakin kecil nilai *error* maka penentuan *cut score* semakin tepat. Selain itu, perlu diperhatikan juga bahwa penentuan *cut score* yang terlalu tinggi dapat menimbulkan kerugian bagi peserta tes. Hal ini dikarenakan *cut score* yang terlalu tinggi menyebabkan peserta tes yang seharusnya lulus menjadi tidak lulus. Sebaliknya jika *cut score* yang terlalu rendah memberikan keuntungan bagi peserta tes. Semakin rendah *cut score* yang ditentukan maka peserta tes yang seharusnya tidak lulus menjadi lulus. Besarnya *cut score* bisa dinaikkan ataupun diturunkan, tetapi kenaikan dan penurunan *cut score* akan berdampak pada besar kecilnya nilai *error*.

Besar kecilnya *error* dari masing-masing *cut score* pada setiap metode *standard setting* menunjukkan ketepatan dari metode tersebut. Namun, penentuan *cut score* pada *standard setting* bukan hanya sebatas melihat besarnya *cut score* dan *errornya*, tetapi untuk menghasilkan *cut score* yang tepat dari metode *standard setting* sebaiknya dilakukan estimasi kesalahan pengukuran dari masing-masing metode *standard setting*.

Sejauh pengamatan peneliti, penelitian mengenai *standard setting* mayoritas melakukan perbandingan metode, tanpa mengestimasi kesalahan dari masing-masing metode. Dengan mengestimasi kesalahan dari setiap metode maka akan diperoleh *cut score* yang lebih akurat. Penelitian yang dilakukan oleh Prijowuntato, Mardapi, & Budiyono (2015, p. 176) merupakan salah satu penelitian yang membahas tentang estimasi kesalahan pengukuran *standard setting*. Penelitian tersebut menggunakan tiga metode *standard setting* yaitu metode *Angoff*, metode *Ebel*, dan

metode *Bookmark*. Estimasi kesalahan pengukuran menggunakan metode *Bootstrap*.

Penelitian serupa dilakukan oleh Yin & Scoring (2008, p. 182) tentang estimasi kesalahan *standard setting* dengan pendekatan *Generalizability Theory*, dimana metode yang digunakan terdiri dari metode *item rating* dan *Bookmark*. Pada penelitian tersebut, *Generalizability Theory* digunakan untuk mengestimasi kesalahan baku *cut score* yang dihasilkan oleh kedua metode *standard setting* yang digunakan yaitu *item rating* dan *Bookmark*. Selanjutnya, *Generalizability Theory* secara eksplisit menggabungkan beberapa sumber kesalahan di model pengukuran yang digunakan untuk mengestimasi kesalahan baku pada *cut score* dari masing-masing metode. Tujuan dalam penelitian tersebut ada tiga yaitu mengestimasi efek dari berbagai sumber kesalahan pada kedua prosedur; mengestimasi standar *error* pada *cut score* dari dua prosedur; dan mengestimasi efek perbedaan konseptualisasi yang berbeda dari seluruh bidang generalisasi untuk dua prosedur *standard setting*.

Berdasarkan latar belakang tersebut, maka estimasi kesalahan pengukuran *cut score* pada beberapa metode *standard setting* perlu dilakukan. Metode *standard setting* yang digunakan dalam penelitian terdiri dari metode *Ebel*, metode *Bookmark*, dan metode *Contrasting group*. Metode *Ebel* dan metode *Bookmark* didasarkan pada tes/butir, dan metode *Contrasting group* didasarkan pada peserta tes (*examinee*). Ketiga metode ini memiliki prosedur yang berbeda dalam penentuan *cut score*. Perbedaan prosedur akan menghasilkan *cut score* dan nilai *error* yang berbeda. *Cut score* yang diperoleh dari ketiga metode ini kemudian diestimasi kesalahannya dengan pendekatan *Generalizability Theory* menggunakan program *eduG*.

Passing scores, cut scores, cut off scores, performance scores, achievement levels, mastery levels, proficiency levels, thresholds levels, dan standard merupakan istilah-istilah dalam standar *setting* (Glass, 1978, p. 240). Istilah-istilah tersebut pada dasarnya memiliki makna yang sama yaitu bahwa *standard setting* merupakan suatu batas atau kriteria yang dijadikan dasar da-

lam suatu hal. *Standard setting* diartikan oleh Cizek (1996, p. 20) sebagai suatu proses dalam menentukan batas lulus (*cut score*). Batas lulus tersebut merupakan batas bawah yang menentukan peserta didik dapat dikatakan kompeten atau tidak kompeten.

MacCann & Stanley (2006, p. 4) mendefinisikan *standard setting* sebagai kumpulan prosedur yang sistematis dalam mengidentifikasi batas lulus (*cut score*) yang diperlukan untuk menentukan tingkat kemahiran. Crocker & Algina, (1986, p. 410) menyebutkan *standard setting* sebagai kegiatan untuk menentukan skor batas lulus. Batas lulus tersebut menjadi kriteria dalam menentukan tingkat atau level prestasi seseorang. *Standard setting* merupakan suatu proses/ prosedur yang sistematis dalam menentukan batas lulus atau *cut score* untuk menyatakan tingkat prestasi.

Pengambilan keputusan berhubungan dengan prosedur-prosedur pengukuran. *Standard setting* merupakan prosedur pengukuran yang dapat digunakan dalam pengambilan keputusan. *Standard setting* adalah aturan yang dapat dipakai dalam pengambilan keputusan yang penting dengan mempertimbangan berbagai aspek.

Standard setting berperan penting dalam bidang pendidikan, tepatnya dalam menentukan batas kelulusan peserta didik. Penyelenggaraan ujian merupakan cara yang digunakan untuk mengevaluasi pembelajaran dengan melihat respon peserta didik terhadap tes yang dibuat. Ujian yang di Indonesia dikenal dengan Ujian Nasional memiliki batas kelulusan atau kriteria kelulusan dengan kata lain peserta tes tersebut dinyatakan kompeten terhadap suatu pelajaran atau materi jika hasil perolehannya melebihi kriteria yang ditentukan. Sebaliknya, seorang peserta tes dikatakan tidak lulus atau tidak kompeten terhadap suatu pelajaran atau materi apabila hasil perolehannya kurang dari kriteria yang ditentukan.

Penelitian ini menggunakan tiga metode *standard setting* yang terdiri dari dua metode berpusat pada tes dan satu metode berpusat pada peserta tes. Dua metode tersebut yaitu metode *Ebel* dan *Bookmark*. Me-

tode yang berpusat pada peserta tes yaitu metode *Contrasting group*.

Metode *Ebel* ini merupakan perbaikan dari metode *Angoff* dengan mempertimbangkan tingkat kesulitan butir dan relevansi butir. Tingkat kesulitan butir pada metode *Ebel* dibedakan menjadi tiga yaitu sulit, sedang, dan mudah. Sementara relevansi butir diperingkat menjadi empat yaitu *essential*, penting (*important*), dapat diterima (*acceptable*), dan dapat dipertanyakan (*questionable*) (Retnawati, 2014, p. 167). Prosedur ini menghasilkan tabel 3 x 4 dengan sejumlah butir tertentu yang diletakkan pada 12 sel kombinasi antara kesulitan butir dan relevansi butir (Alsmadi, 2007, p. 479).

Prosedur *Ebel* dapat dilakukan dengan penilai menentukan tingkat kesulitan butir (sulit, sedang, mudah) dan relevansi butir meliputi *essential*, penting, dapat diterima, dan dapat dipertanyakan (Saunders, Ryan, & Huynh, 1980, p. 167); penilai mengisi butir-butir pada sel kombinasi tingkat kesulitan dan relevansi butir; penilai menentukan proporsi butir dalam masing-masing kategori bahwa *examinee* yang ada pada garis batas; penilai mengalikan jumlah butir dengan proporsi masing-masing. Hasil perkalian tiap kategori tersebut kemudian dijumlahkan. Hasil penjumlahan ini disebut *minimum passing score* (MPS); *cut score* diperoleh dengan merata-rata *minimum passing score* yang diusulkan oleh penilai.

Metode *Bookmark* dikembangkan untuk mengatasi keterbatasan yang berhubungan dengan *standard setting* yang terdahulu, khususnya untuk menentukan *cut score* ganda pada *single test* (Karantonis & Sireci, 2006, p. 6). Pengembangan metode ini mencakup beberapa prosedur seperti mengintegrasikan *selected response* dan *constructed response* ke dalam format butir, mudah diterapkan, dan mendasarkan pada IRT (Cizek & Bunch, 2007, p. 160). Pada metode *Bookmark*, butir-butir yang sudah diurutkan dengan menggunakan analisis IRT dikumpulkan dalam satu booklet yang disebut *Ordered Item Booklet* (OIB). Di samping itu, dalam metode *Bookmark* ditetapkan *Response Probability* (RP) sebesar 67% *likelihood*.

Adapun kelemahan implementasi prosedur *Bookmark* adalah komposisi soal dari yang termudah sampai paling sulit memungkinkan ada beberapa indikator dari kemampuan yang diujikan tidak termasuk dalam halaman *Bookmark* yang dipilih oleh panelis sebagai batas kemampuan siswa dalam menjawab. Disamping itu, kelemahan metode *Bookmark* adalah penilai kesulitan untuk memahami dan menggunakan kemungkinan jawaban (*Response Probability*).

Metode *Contrasting group* diperkenalkan oleh Berk pada tahun 1976. Berk menyarankan prosedur validasi kelompok yang diperluas. Prosedur kelompok yang menguasai digunakan untuk menentukan perbedaan *cut score* antara siswa-siswa yang terlatih dan tak terlatih, atau antara siswa yang menguasai materi dan tidak menguasai materi (Cizek & Bunch, 2007, p. 106). Penentuan *cut score* pada metode ini menggunakan prosedur kelompok untuk membedakan kelompok master dan kelompok non-master. Dua distribusi kelompok tersebut kemudian digambar untuk menentukan titik potongnya.

Metode *Contrasting group* memiliki kelebihan dan kekurangan. Livingstone & Zieky (1982, p. 53) menganggap bahwa metode ini memiliki kemudahan dalam penerapannya dan memberikan hasil yang akurat. Metode ini didasarkan pada kondisi nyata peserta tes. Selain itu, tes dengan bentuk pilihan ganda cocok jika penentuan *cut score* dilakukan dengan metode ini, karena ahli akan menentukan kelompok master dan non-master dengan lebih mendasar. Sementara kekurangan metode ini adalah adanya kesulitan dalam memperoleh evaluasi yang sebanding untuk wilayah yang lebih luas, misal tingkat nasional.

Masalah dalam *standard setting* pada dasarnya sama dengan masalah yang dihadapi dalam pengukuran (Nichols, Twing, Mueller, & O'Malley, 2010, p. 19). Tidak tersedianya indikator dalam *standard setting* yang dapat digunakan untuk mengukur prestasi kelompok. Para panelis dalam *standard setting* diminta untuk membuat pertimbangan tentang kinerja kelompok siswa, sebagai

contoh siswa yang termasuk dalam kategori dasar, cukup, maupun maju.

Panelis yang terlibat dalam putaran dan skema dapat menyebabkan variabilitas dalam *cut score* (Yin & Sconing, 2008, p. 185). Kemungkinan besar bahwa perbedaan *cut score* dapat dihubungkan dengan perbedaan metode *standard setting* karena berbagai sumber yang disebutkan di atas dan atau perbedaan dalam prosedur. *Standard setting* termasuk prosedur pengukuran untuk menetapkan kemampuan siswa, maka variabilitas atau ketidakpastian dalam *cut score* yang dihasilkan dari proses *standard setting* perlu diperhatikan. *Standard error* dalam *cut score* seharusnya ikut dipertimbangkan. Namun demikian, sedikit penelitian yang mempertimbangkan *standard error* dalam *standard setting*.

Estimasi kesalahan pengukuran *cut score* dalam *standard setting* pada penelitian ini dilakukan dengan pendekatan *Generalizability Theory*. Pendekatan *generalizability theory* digunakan karena dalam metode *standard setting* ini digunakan panelis sebagai penentu *cut score*, dimana penggunaan panelis akan menimbulkan tingginya variabilitas nilai yang diberikan oleh panelis itu sendiri. Sehingga reliabilitas dari skor yang dihasilkan perlu diperhatikan. Penggunaan *generalizability theory* dikarenakan *G theory* memberikan berbagai model yang digunakan dalam menyelidiki kesalahan dalam metode *Ebel*, *Bookmark*, dan *Contrasting group*.

Analisis G-teori memiliki dua tahap yaitu generalisasi studi (G studi) dan *decision* studi (D studi). G studi dilakukan untuk menentukan seberapa baik skor dapat digunakan dalam beberapa situasi dan melibatkan perkiraan komponen varians yang mungkin akan digunakan dalam studi D. Studi D adalah penelitian yang dilakukan untuk menghitung koefisien reliabilitas dan SEs pengukuran dengan tujuan menentukan prosedur pengukuran yang paling efisien pada situasi tertentu. Pertimbangan studi D yang paling penting adalah spesifikasi dari generalisasi populasi dimana pembuat keputusan ingin menggeneralisasi skor dengan prosedur pengukuran tertentu.

Berdasarkan uraian yang telah disampaikan tersebut, maka penelitian ini bertujuan untuk mengetahui metode yang lebih akurat dalam mengestimasi kesalahan pengukuran *standard setting* pada metode *Ebel*, *Bookmark*, dan *Contrasting group*.

Metode

Penelitian ini termasuk jenis penelitian deskriptif kuantitatif. Sumber data dalam penelitian ini berupa respon siswa terhadap Ujian Nasional mata pelajaran Matematika pada jenjang SMP di Kabupaten Sumbawa tahun ajaran 2015/2016

Populasi dalam penelitian ini adalah lembar jawaban Ujian Nasional Matematika siswa dari 95 SMP Negeri di Kabupaten Sumbawa. Adapun sampel yang digunakan dalam penelitian ini merupakan respon jawaban Ujian Nasional Matematika siswa paket P0C5520 dari 12 SMP Negeri di Kabupaten Sumbawa dengan jumlah sampel sebanyak 352 siswa. Pengambilan sampel dalam penelitian ini menggunakan teknik *proportionate stratified random sampling*. Pengambilan sampel didasarkan pada klasifikasi sekolah dari kategori tinggi, sedang, dan rendah. Klasifikasi sekolah berdasarkan pada nilai Ujian Nasional tahun ajaran 2015/2016. Selain itu juga mempertimbangkan letak geografis sekolah baik itu di kota maupun di desa. Hal ini dilakukan agar sampel yang digunakan dalam penelitian dapat merepresentasikan keadaan sebenarnya di Kabupaten Sumbawa. Data tentang sampel penelitian disajikan pada Tabel 1.

Tabel 1. Daftar Sampel Penelitian

No	Nama Sekolah	Kategori	Letak	Jumlah Sampel
1	SMP N 1 Sumbawa Besar	Tinggi	Kota	58
2	SMP N 1 Moyo Hilir	Tinggi	Desa	21
3	SMP N 1 Moyo Utara	Tinggi	Desa	25
4	SMP N 5 Moyo Hilir	Tinggi	Desa	10
5	SMP N 2 Labuhan Badas	Rendah	Desa	15
6	SMP N 2 Sumbawa Besar	Rendah	Kota	62
7	SMP N 3 Moyo Hilir	Rendah	Desa	10
8	SMP N 3 Sumbawa Besar	Rendah	Kota	32
9	SMP N 4 Labuhan Badas	Rendah	Desa	17
10	SMP N4 Sumbawa Besar	Rendah	Kota	7
11	SMP N 1 Unter Iwes	Rendah	Kota	34
12	SMP N 1 Labuhan Badas	Rendah	Desa	61
Total				352

Panelis yang digunakan ditentukan dengan kuota sebanyak 12 orang guru matematika yang diambil berdasarkan kualitas sekolah (tinggi, sedang, dan rendah) dan letak geografisnya (kota dan desa) serta kriteria panelis. Adapun kriteria tersebut sebagai berikut: (1) ahli dalam bidang yang berhubungan dengan ujian; (2) terbiasa dengan metode-metode ujian; (3) telah mengajar matematika minimal 10 tahun; (4) mengajar matematika kelas 12 minimal selama 5 tahun; dan (5) lulusan dari program studi matematika atau pendidikan matematika.

Teknik yang digunakan untuk mengumpulkan data dalam penelitian ini adalah dokumentasi. Teknik dokumentasi digunakan untuk mengumpulkan respon jawaban siswa SMP di Kabupaten Sumbawa dalam menjawab Ujian Nasional mata pelajaran Matematika tahun ajaran 2015/2016. Selain itu, data dalam penelitian ini juga dikumpulkan melalui *Focus Group Discussion* (FGD). Adapun instrumen yang digunakan untuk mengumpulkan data dari FGD yaitu berupa lembar kerja panelis untuk metode *Ebel*, *Bookmark*, dan *Contrasting group*.

Teknik analisis data dalam penelitian ini terdiri dari tiga tahap yaitu tahap pertama atau persiapan terdiri dari penyiapan data, penggolongan SMP, dan pengujian karakteristik butir menggunakan program *Winstep*; tahap kedua yaitu *focus group discussion* (FGD) terdiri dari dua putaran; dan tahap ketiga yaitu mengestimasi kesalahan pengukuran dengan pendekatan *Generalizability Theory* menggunakan program *EduG*.

Hasil dan Pembahasan

Data respon Ujian Nasional Matematika siswa paket P0C5520 dari 12 SMP Negeri di Kabupaten Sumbawa dengan jumlah sampel sebanyak 352 siswa dengan soal terdiri dari 40 butir. Selanjutnya dianalisis menggunakan program *Winstep* untuk menghitung tingkat kesulitan butir. Tingkat kesulitan butir pada *output Winsteps* dapat dilihat pada *Table Measure*. Hasil analisis *Winsteps* tampak pada Tabel 2.

Tabel 2. Tingkat Kesulitan Butir

No	Kriteria	Nomor Butir	Jumlah
1	Mudah	1, 10, 16, 39	4
2	Sedang	4, 5, 6, 7, 8, 9, 11, 12, 13, 14, 15, 18, 20, 21, 22, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 37, 38, 40	30
3	Sulit	2, 3, 17, 19, 23, 36	6
Jumlah			40

Berdasarkan Tabel 2 diketahui bahwa sebanyak 4 butir (10%) tergolong butir yang mudah, 30 butir (75%) termasuk butir yang tergolong sedang atau yang biasa disebut sebagai butir yang baik, dan butir yang termasuk dalam kategori butir sulit berjumlah 6 butir (15%).

Tahap kedua yaitu FGD untuk menentukan *cut score* dari metode *Ebel*, *Bookmark*, dan *Contrasting group*. Tahap ini dilakukan dengan dua putaran. Masing-masing metode memiliki dua *cut score*. Berikut *cut score* tiap metode disajikan pada Tabel 3.

Tabel 3. *Cut Score* Putaran Pertama dan Kedua

Putaran	Metode <i>Standard Setting</i>		
	Ebel	Bookmark	Contrasting Group
Pertama	64,579	64,434	50
Kedua	67,813	65,455	47,5

Ketiga metode menghasilkan *cut score* yang berbeda. Perbedaan ini terjadi dikarenakan prosedur penentuan *cut score* masing-masing metode sangat berbeda. Metode *Ebel* memiliki *cut score* tertinggi pada kedua putaran yaitu 64,579 dan 67,813 dengan jumlah siswa yang lulus sebanyak 128 siswa dan mampu menguasai minimal 26 deskriptor. *Cut score* metode *Bookmark* putaran 1 dan 2 berturut 64,434 dan 65,455. Jumlah siswa yang lulus pada *cut score* metode *Bookmark* sebanyak 128 siswa dengan deskriptor yang harus dikuasai sebanyak 25 deskriptor. Metode *Contrasting group* memiliki *cut score* paling rendah pada kedua putaran dengan jumlah siswa yang lulus sebanyak 297 siswa. Adapun siswa dikatakan lulus jika mampu menguasai 19 deskriptor.

Rendahnya kelulusan siswa tersebut dapat disebabkan oleh kekurangan telitian peserta didik dalam mengerjakan soal. Walaupun soal yang diujikan termasuk dalam kriteria sedang, namun apabila peserta didik tidak teliti dalam mengerjakan maka hasil pengerjaan tetap salah. Padahal berdasarkan pendapat panelis soal-soal yang diujikan termasuk dalam kategori esensial dan penting, yang mana berarti materi-materi yang berkaitan dengan soal tersebut harusnya dapat dipahami dan dimengerti oleh peserta didik.

Berdasarkan hasil penelitian ini penyebab tingginya *cut score* pada metode *Ebel* dikarenakan masih kurangnya pemahaman guru tentang relevansi butir dengan kategori esensial, penting, dapat diterima, dan dapat dipertanyakan serta tingkat kesulitan butir. Selain itu, perkiraan proporsi jawaban benar untuk masing-masing kategori juga menjadi penyebab tingginya *cut score*. Kurangnya pemahaman guru akan hal-hal penting pada metode *Ebel* membuat guru dengan bebas meletakkan soal pada keempat kategori dengan proporsi jawaban benar yang tinggi sesuai dengan kemampuan guru dalam menginterpretasi soal.

Prosedur *Bookmark* merupakan prosedur yang dirasa mudah bagi panelis dalam menetapkan *cut score*. Hal ini dikarenakan data yang digunakan untuk menentukan *cut score* sudah disiapkan terlebih dahulu. Sehingga dapat mempermudah panelis dalam menentukan halaman *bookmark* yang dianggap sesuai dengan kemampuan peserta didik menjawab benar.

Berbeda dengan metode *Ebel* dan *Bookmark*, pada metode metode *Contrasting group* putaran pertama dan kedua diperoleh *cut score* yang lebih rendah dari kedua metode tersebut. Karakteristik parameter butir soal secara tidak langsung akan mempengaruhi hasil perhitungan *cut score* yang diperoleh dengan menggunakan metode *Contrasting Group*. Jika butir-butir soal yang terdapat dalam perangkat tes memiliki tingkat kesulitan rendah maka butir soal akan dijawab dengan benar sehingga skor peserta tes menjadi tinggi. Sebaliknya jika butir-butir soal yang terdapat dalam perangkat tes memiliki

tingkat kesulitan tinggi maka butir soal akan susah dijawab dengan benar dan menyebabkan skor peserta tes menjadi rendah. Distribusi skor inilah yang mempengaruhi *cut score* yang diperoleh.

Setelah diperoleh *cut score* pada ketiga metode *standard setting*, selanjutnya dilakukan tahap ketiga. Tahap ketiga yaitu mengestimasi kesalahan pengukuran dengan pendekatan *generalizability theory*. Pendekatan ini dilakukan dengan dua tahap yaitu G studi dan D studi.

G studi dilakukan untuk menentukan seberapa baik skor dapat digunakan dan melibatkan perkiraan varian komponen yang akan digunakan dalam studi D. D studi dilakukan untuk menghitung koefisien reliabilitas dan SEs pengukuran dengan tujuan menentukan prosedur pengukuran yang paling efisien. Analisis pada pendekatan ini dilakukan dengan program *EduG*. Hasil analisis dapat dilihat pada Tabel 4.

Tabel 4. Estimasi Varians Komponen dengan Desain G studi

VC	Ebel		Bookmark		Contrasting group	
	Estimasi	%	Estimasi	%	Estimasi	%
	31,41	76,8	5,35	77,4	0,00	0,0
	5,08	13,6	0,39	2,4	3,12	100
	1,79	9,6	1,60	20,1	0,00	0,0

Hasil menunjukkan bahwa estimasi varians komponen putaran terbesar pada metode *Ebel*, *Contrasting group*, dan *Bookmark*. Namun dilihat dari proporsinya varian komponen putaran pada metode *Contrasting group* lebih tinggi dibanding metode *Ebel* dan *Bookmark*. Tingginya proporsi varians komponen putaran pada metode *Contrasting group* dikarenakan pada metode ini hanya terdapat dua *cut score* yaitu putaran satu dan dua. Dimana *cut score* setiap panelis untuk masing-masing putaran itu sama. Sehingga menyebabkan variabilitas *cut score* hanya tinggi pada putaran. Variabilitas inilah yang akan berpengaruh pada koefisien reliabilitas.

Setiap G studi akan diperoleh nilai koefisien G *relative* dan G *absolute* pada masing-masing metode. Koefisien G diguna-

kan untuk menentukan besarnya koefisien reliabilitas setiap metode atau dengan kata lain koefisien G disebut sebagai koefisien reliabilitas. Adapun koefisien G dapat dilihat pada Tabel 5.

Tabel 5. Koefisien G pada Metode *Ebel*, *Bookmark* dan *Contrasting group*

Koefisien	Ebel	Bookmark	Contrasting group
G relative	0,94	0,59	1,00
G Absolute	0,65	0,23	1,00

Tabel 5 menunjukkan bahwa koefisien G tertinggi dari ketiga metode tersebut yaitu diperoleh pada metode *Contrasting group* sebesar 1,00. Koefisien G terendah diperoleh oleh metode *Bookmark*. Hal ini berarti bahwa metode *Contrasting group* memiliki koefisien reliabilitas tertinggi atau dengan kata lain metode *Contrasting group* lebih reliabel daripada metode *Ebel* dan *Bookmark*. Selanjutnya dilakukan D studi untuk melihat reliabilitas dari masing-masing optimalisasi pada setiap metode.

Tahap kedua pada pendekatan *Generalizability Theory* yaitu D studi. D studi dilakukan untuk menghitung koefisien reliabilitas dan *standard error* pada masing-masing metode. *Standard error* yang digunakan berdasarkan hasil analisis dengan eduG dapat dilihat pada *absolute standard error of measurement* (SEM). Masing-masing metode *standard setting* terdiri dari lima SEM yang berbeda. Hasil SEM pada *cut score* masing-masing metode dapat dilihat pada Tabel 6.

Tabel 6. *Standard Error of Measurement* pada *Cut Score*

SEM	Ebel	Bookmark	Contrasting Group
		(1)	
	1,158	0,803	0,000
		(2)	
	1,113	0,772	0,000
		(3)	
	1,072	0,744	0,000
		(4)	
	1,036	0,718	0,000
		(5)	
	1,036	0,696	0,000

Berdasarkan hasil tersebut terlihat bahwa semakin banyak jumlah panelis yang digunakan dengan jumlah putaran yang sama maka nilai SEM semakin rendah. Nilai SEM tertinggi pada metode *Ebel* yang mencapai 1,158 dengan jumlah panelis sebanyak 12 orang. Nilai SEM terendah pada metode *Contrasting group* sebesar 0,000 untuk semua pola optimalisasi.

Estimasi varians komponen pada ketiga metode *standard setting* menunjukkan metode *Contrasting group* memiliki nilai estimasi terendah pada efek panelis dan interaksi panelis dan putaran. Estimasi varian komponen pada efek dan interaksi panelis dan putaran tertinggi pada metode *Ebel*. Estimasi VC tertinggi menunjukkan tingginya variabilitas nilai yang diberikan panelis dan sumber lain, seperti tingkat kesulitan butir soal dan tingkat penguasaan peserta didik yang membutuhkan *judgment* panelis. Hal ini senada dengan yang disampaikan oleh Lee & Lewis (2008, p. 614).

Hasil varian komponen yang didasarkan pada G studi menjadi dasar dilakukannya D studi. Hasil D studi digunakan untuk menentukan koefisien reliabilitas, SEM dan keterkaitan *universe* untuk bisa digeneralisasi. Perbandingan nilai SEM untuk menentukan metode yang paling akurat dalam menentukan *cut score*. Nilai SEM terkecil diantara ketiga metode merupakan metode paling akurat.

Seperti yang diharapkan, *universe* pada bentuk random menghasilkan SE yang lebih besar karena *universe* memiliki definisi yang lebih luas untuk digeneralisasi daripada *universe* bentuk *fixed*. *Universe* bentuk *fixed* lebih dibatasi, akibatnya menghasilkan SE yang lebih kecil. Secara umum, SE *cut score* metode *Contrasting group* relatif lebih kecil daripada SE *cut score* metode *Ebel* dan *Bookmark*.

SEM pada ketiga metode dengan jumlah panelis dua belas dan putaran sebanyak dua kali. Hal ini menunjukkan bahwa nilai SEM pada metode *Contrasting group* lebih kecil dari metode *Ebel* dan metode *Bookmark*. Berarti bahwa metode *Contrasting group* lebih akurat daripada dua metode lainnya.

Rendahnya nilai SEM pada metode *Contrasting group* dikarenakan prosedur metode ini tidak membutuhkan penilaian dari setiap panelis, yang mana panelis melakukan diskusi untuk menentukan kelompok master dan nonmaster berdasarkan pengetahuan dan pengalaman panelis selama mengajar. Penentuan kelompok master dan nonmaster dilakukan hanya pada pembagian sekolah. Selanjutnya para panelis secara bersama menentukan jumlah peserta didik yang mampu mengerjakan butir dengan benar berdasarkan interval yang digunakan dalam penelitian ini. Kemudian terbentuklah distribusi frekuensi yang dijadikan sebagai penentu *cut score*. Sehingga hanya dihasilkan satu *cut score* dari dua belas panelis tanpa adanya istilah merata-ratakan *cut score* panelis.

Prosedur pada metode *Contrasting group* sangat berbeda dengan dua metode lainnya. Dimana metode *Ebel* dan *Bookmark* menghasilkan dua belas *cut score* yang selanjutnya dirata-ratakan untuk dijadikan sebagai *cut score* metode tersebut. Selain itu, metode *Contrasting group* merupakan metode yang paling mudah untuk diterapkan dibandingkan metode *Ebel* dan *Bookmark*. Sehingga, menyebabkan metode *Ebel* dan *Bookmark* memiliki variabilitas nilai yang tinggi.

Selain itu, koefisien G dari ketiga metode menunjukkan bahwa metode *Contrasting group* memiliki koefisien G yang lebih besar dari metode *Ebel* dan *Bookmark*. Hal ini berarti bahwa metode *Contrasting group* lebih reliabel dibandingkan metode *Ebel* dan *Bookmark*. Dikarenakan variabilitas *cut score* pada metode *Contrasting group* rendah.

Simpulan

Berdasarkan hasil penelitian yang telah dilakukan terkait dengan perbandingan estimasi kesalahan pengukuran *standard setting* pada penilaian kompetensi Matematika SMP di Kabupaten Sumbawa dapat disimpulkan bahwa *cut score* mata pelajaran Matematika Jenjang SMP di Kabupaten Sumbawa yang dihasilkan dengan menggunakan metode *Ebel* pada putaran pertama adalah 64,579 mengalami kenaikan pada putaran kedua menjadi 67,813. Pada metode ini siswa

dikatakan lulus jika mampu menguasai 26 deskriptor. Selanjutnya pada metode *Bookmark* diperoleh *cut score* putaran pertama sebesar 64,434 mengalami penurunan pada putaran kedua menjadi 65,455, dimana siswa dikatakan lulus jika menguasai 25 deskriptor.

Cut score mata pelajaran Matematika jenjang SMP di Kabupaten Sumbawa yang dihasilkan dengan menggunakan metode *Contrasting group* putaran 1 sebesar 50,00 mengalami penurunan pada putaran kedua menjadi 47,5. Terdapat 19 deskriptor kemampuan Matematika yang harus dikuasai oleh siswa agar dapat lulus berdasarkan metode *Contrasting group*. Metode *Contrasting group* merupakan metode yang paling akurat untuk mengestimasi kesalahan pengukuran dibandingkan metode *Ebel* dan *Bookmark*. Hal ini dikarenakan variabilitas *cut score* pada metode *Contrasting group* rendah, memiliki nilai SEM yang rendah, dan koefisien G yang tinggi.

Saran yang dapat diberikan berdasarkan simpulan penelitian ini adalah sebagai berikut. Bagi Dinas Pendidikan Kabupaten Sumbawa, perlu adanya pelatihan mengenai penentuan *cut score* bagi kelompok guru mata pelajaran sehingga guru memiliki tambahan pengetahuan yang lebih dan akhirnya mampu menerapkan metode-metode tersebut dalam menentukan *cut score*; bagi peneliti berikutnya, perlu dilakukan penelitian lebih lanjut terkait dengan *cut score* tentang penguasaan materi matematika di sekolah menengah pertama; dan bagi peneliti berikutnya, perlu dilakukan penelitian lebih lanjut terkait dengan *cut score* tentang penguasaan materi matematika di sekolah menengah pertama.

Daftar Pustaka

- Alsmadi, A. A. (2007). A comparative study of two standard-setting technique. *Social Behavior and Personality*, 38(4), 479–486.
- Berk, R. A. (1986). A Consumer's guide to setting performance standards on criterion-referenced tests. *Review of Educational Research*, 56(1), 137–172.

- <https://doi.org/10.3102/00346543056001137>
- Cizek, G. J. (1996). An NCME instructional module on: setting passing scores. *Educational Measurement: Issues and Practice*, 15(2), 20–31. <https://doi.org/10.1111/j.1745-3992.1996.tb00809.x>
- Cizek, G. J., & Bunch, M. B. (2007). *Standard setting: a guide to establishing and evaluating performance standards for tests*. California: Sage Publication, Inc.
- Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory*. New York: CBS College Publishing.
- Glass, G. V. (1978). Standards and criteria. *Journal of Educational Measurement*, 15(4), 237–261. <https://doi.org/10.1111/j.1745-3984.1978.tb00072.x>
- Karantonis, A., & Sireci, S. G. (2006). The bookmark standard-setting method: a literature review. *Educational Measurement: Issues and Practice*, 25(1), 4–12. <https://doi.org/10.1111/j.1745-3992.2006.00047.x>
- Lee, G., & Lewis, D. M. (2008). A Generalizability theory approach to standard error estimates for bookmark standard settings. *Educational and Psychological Measurement*, 68(4), 603–620. <https://doi.org/10.1177/0013164407312603>
- Livingstone, S. A., & Zieky, M. J. (1982). *Passing scores: a manual for setting standards of performance on educational and occupational tests*. Princeton, New Jersey: Educational Testing Service.
- MacCann, R. G., & Stanley, G. (2006). The use of rasch modeling to improve standard setting. *Practical Assessment, Research & Evaluation*, 11(2), 1 – 17.
- Menteri Pendidikan dan Kebudayaan Republik Indonesia. Peraturan Menteri Pendidikan dan Kebudayaan Republik Indonesia Nomor 5 Tahun 2015 tentang Kriteria Kelulusan Peserta Didik, Penyelenggaraan Ujian Nasional, dan Penyelenggaraan Ujian Sekolah/Madrasah/Pendidikan Kesetaraan Pada Smp/Mts atau yang Sederajat d (2015).
- Nichols, P., Twing, J., Mueller, C. D., & O'Malley, K. (2010). Standard-setting methods as measurement processes. *Educational Measurement: Issues and Practice*, 29(1), 14–24. <https://doi.org/10.1111/j.1745-3992.2009.00166.x>
- Nudell, H. (2008). Making the cut score, that is establishing a pass/fail score is a highly technical process. *ICSC Certified Professionals Newsletter*.
- Prijowuntato, S. W., Mardapi, D., & Budiyo, B. (2015). Perbandingan estimasi kesalahan pengukuran standard setting dalam penilaian kompetensi akuntansi SMK. *Jurnal Penelitian Dan Evaluasi Pendidikan*, 19(2). Retrieved from <https://journal.uny.ac.id/index.php/jpep/article/view/5578>
- Retnawati, H. (2014). *Teori respon butir dan penerapannya*. Yogyakarta: Nuha Medika.
- Saunders, J. C., Ryan, J. P., & Huynh, H. (1980). *A comparison of two ways of setting passing scores based on the nedelsky procedure*. Publication Series in Mastery Testing. South Carolina: University of South Carolina.
- Yin, P., & Sconing, J. (2008). Estimating standard errors of cut scores for item rating and mapmark procedure: a generalizability theory approach. *Educational and Psychological Measurement*, 68(1), 182–197.
- Zieky, M. J., Perie, M., & Livingston, S. A. (2008). *Cutscores: a manual for setting standards of performance on educational and occupational tests*. Princeton, New Jersey: Educational Testing Service.

EVALUASI PROGRAM DIKLAT KARYA TULIS ILMIAH UNTUK WIDYAISWARA PUSBANGTENDIK KEMDIKBUD

Riyan Arthur

Fakultas Teknik, Universitas Negeri Jakarta

Rawamangun, Pulogadung, Kota Jakarta Timur, DKI Jakarta 13220, Indonesia

Email: arthur@unj.ac.id

Abstrak

Tujuan penelitian ini adalah mengevaluasi program Pendidikan dan Pelatihan Karya Tulis Ilmiah (Diklat KTI) untuk widyaiswara yang diselenggarakan Pusbangtendik Kemdikbud. Penelitian ini menggunakan metode evaluatif, model Kirkpatrick. Penelitian ini melibatkan 32 widyaiswara. Pengumpulan data dilakukan dengan menyebarkan kuesioner, observasi, studi dokumentasi dan wawancara. Hasil penelitian menunjukkan, Pertama; hanya ada tiga level Diklat KTI yang dilaksanakan sesuai peraturan, perencanaan serta memenuhi kriteria evaluasi, yaitu: level reaksi, pembelajaran dan perilaku. Kedua; Diklat KTI yang diselenggarakan Pusbangtendik Kemdikbud masih belum berdampak terhadap individu dan unit kerja. Rekomendasi atas Program Diklat KTI adalah dilanjutkan dengan perbaikan pada level dampak Diklat KTI.

Kata kunci: *evaluasi, Diklat KTI, widyaiswara*

THE EVALUATION OF SCIENTIFIC PAPER TRAINING PROGRAM FOR PUSBANGTENDIK KEMDIKBUD TRAINERS

Abstract

This research is aimed at evaluating the programs of education and training of scientific papers (KTI training) for trainers in Pusbangtendik Kemdikbud. This research used an evaluative method of Kirkpatrick model. This study involved 32 trainers. Data were collected by distributing questionnaires, observation, documentation study, and interview. The result shows that, first, there are only three levels of KTI Training implemented according to the regulation, plans and fulfillment of the evaluation criteria; the three levels are reaction, learning, and behaviour levels. Second, KTI Training held by Pusbangtendik Kemdikbud still has not significantly showed any impact towards the individual and work units. Therefore, it is recommended for KTI Training program to continue with the improvement efforts on the impact level of KTI Training.

Kata kunci: *evaluation, KTI training, trainer*

Permalink/DOI: <http://dx.doi.org/10.21831/pep.v22i1.16749>

Pendahuluan

Dalam upaya meningkatkan kualitas widyaiswara, Pusat Pengembangan Tenaga Kependidikan (Pusbangtendik) Kementerian Pendidikan dan Kebudayaan (Kemdikbud) bekerja sama dengan Lembaga Administrasi Negara (LAN) pada tahun 2015 menyelenggarakan Pendidikan dan Pelatihan Penyusunan Karya Tulis Ilmiah (Diklat KTI) untuk widyaiswara pertama/muda. Diklat ini berisikan materi tentang prosedur dan metode ilmiah serta kaidah tata tulis yang lazim berlaku dalam komunitas ilmiah agar pembaca KTI dapat memahami secara tepat jalan pikiran dan kandungan materi yang termuat dalam KTI widyaiswara yang bersangkutan.

Diklat kewidyaiswaraan ini memiliki keluaran (*outcomes*) berupa publikasi ilmiah yang dilakukan oleh widyaiswara itu sendiri, baik melalui seminar maupun jurnal. Hal ini merujuk pada Peraturan Kepala Lembaga Administrasi Negara (PERKALAN) nomor 9 tahun 2008 tentang Pedoman penyusunan karya Tulis Ilmiah bagi widyaiswara.

Kaitannya dengan widyaiswara Kemdikbud ditemukan bahwa publikasi ilmiah yang dilakukan widyaiswara tingkat nasional (kecuali pulau Jawa) hanya mencapai skor 6,3 untuk pulau Sumatera, 5,5 untuk Kalimantan, 7,2 untuk Bali dan Nusa Tenggara

Barat, 2,9 untuk Maluku dan Papua dari rentang skor 1-10 (Anisah, Sukmawati, & Made, 2015, p. 9). Selengkapnya dapat dilihat pada Tabel 1.

Uraian pada Tabel 1 menggambarkan bahwa kemampuan widyaiswara dalam menyusun KTI masih lemah. Kelemahan tersebut bukan hanya dari segi penulisan tapi juga pada minimnya publikasi. Padahal tugas pokok dan fungsi jabatan fungsional Widyaiswara dalam PERMENPAN No. 22 tahun 2014 tentang Jabatan Fungsional Widyaiswara salah satunya adalah menulis KTI.

Lebih lanjut dijelaskan bahwa kondisi objektif widyaiswara selama ini masih didominasi oleh kegiatan Dikjartih (mendidik mengajar dan melatih) dan belum banyak melakukan kegiatan pengembangan profesi khususnya penulisan karya tulis ilmiah. Pengembangan motivasi widyaiswara diantaranya dilakukan dengan langkah-langkah: (a) melakukan peningkatan pemahaman dan penguasaan penulisan karya tulis ilmiah melalui diklat penulisan karya tulis ilmiah, (b) menggali topik-topik permasalahan yang akan dikembangkan dalam bentuk karya tulis ilmiah, (c) perlunya melakukan pengembangan *knowledge sharing* forum yang melakukan konsolidasi keilmuan antar Ikatan Peneliti LAN dengan widyaiswara LAN. (Alie, 2015, p. 96).

Tabel 1. Kinerja Peneliti Widyaiswara* (2012-2014)

No	Kinerja peneliti (Widyaiswara Nasional)	Sumatera	Kalimantan	Bali dan NTB	Sulawesi	Maluku dan Papua
1	KTI yang tidak diterbitkan (laporan kegiatan penelitian)	2,4	2,8	4	2,8	2,3
2	KTI hasil penelitian dan pengembangan atau tinjauan/ulasan, tidak/belum diterbitkan, dan disampaikan dalam pertemuan ilmiah	1,8	1,6	1,3	2,1	1,4
3	KTI terbit dalam prosiding pertemuan ilmiah nasional	3,9	2,5	4,3	3,5	0,6
4	KTI terbit dalam majalah ilmiah nasional tidak terakreditasi	1,4	1	1,6	1	1,9
5	KTI terbit dalam majalah ilmiah nasional terakreditasi	0,7	1,6	1	1,2	0,3
6	KTI terbit dalam bentuk bagian dari buku, penerbit nasional	0,3	0,4	0,3	0,6	0,1
Skor publikasi		6,3	5,5	7,2	6,3	2,9

Sumber: (Anisah et al., 2015). *skala 0-10

Hasil kajian di atas, menegaskan bahwa widyaiswara masih belum menyadari sepenuhnya tentang manfaat KTI. Diklat KTI yang diadakan Kemdikbud merupakan salah satu sarana yang dapat digunakan untuk meningkatkan kemampuan widyaiswara dalam menyusun KTI.

Lebih lanjut, ditegaskan Hidayat & Sa'ud (2015, p. 31) bahwa Kompetensi widyaiswara rata-rata *miss-match* atau tidak sesuai dengan bidang studi yang diampu dalam pelaksanaan pendidikan dan pelatihan, Widyaiswara belum sepenuhnya memahami tugas pokok dan fungsinya, Profil kompetensi widyaiswara belum sesuai dengan standar kompetensi widyaiswara yang berlaku, Badan Diklat belum sepenuhnya melaksanakan program pengembangan dan peningkatan kompetensi widyaiswara secara ter-program.

Jika dilihat data-data dan uraian tersebut di atas terlihat pelaksanaan diklat utamanya Diklat KTI selama ini masih belum memberikan dampak yang baik bagi widyaiswara. Oleh karena itu, penting dilakukan evaluasi terhadap program Diklat KTI.

Berdasarkan uraian yang telah disampaikan tersebut maka tujuan penelitian ini adalah mengevaluasi program Pendidikan dan Pelatihan Karya Tulis Ilmiah (Diklat KTI) untuk widyaiswara yang diselenggarakan Pusbangtendik Kemdikbud.

Metode Penelitian

Penelitian ini dilakukan di Pusbangtendik Kemdikbud Depok, Jawa Barat. Penelitian ini dilakukan pada Maret 2015 sampai Januari 2016.

Subjek penelitian ini adalah widyaiswara yang berasal dari instansi-instansi di bawah naungan Kemdikbud. Pemilihan responden ditentukan dengan cara *purposive sampling*. Adapun kriteria dan pertimbangan pemilihan sampel didasarkan pada kriteria Diklat KTI, yaitu: pegawai Negeri Sipil (PNS) yang mempunyai jabatan fungsional widyaiswara pertama dan widyaiswara muda di lingkungan Kemdikbud dimana baru pertama kali mengikuti Diklat KTI.

Metode penelitian yang digunakan dalam penelitian ini adalah metode evaluatif.

Menurut Arikunto & Jabar (2014, p. 59) penelitian evaluasi (evaluatif) adalah penelitian yang mengetahui akhir dari sebuah kebijakan, dalam rangka menentukan rekomendasi atas kebijakan yang lalu, yang pada tujuan akhirnya adalah untuk menentukan kebijakan selanjutnya. *Program evaluation aims to know the achievement of program objectives that have been implemented. Furthermore, the results of the program evaluation are used as the basis for carrying out follow-up activities or for subsequent decision making* (Arthur, 2015, pp. 964–965)

Model yang digunakan dalam penelitian ini adalah model Kirkpatrick. Model Kirkpatrick disebut juga empat level model evaluasi (*four levels of evaluation model*). Model ini dimulai dari level reaksi (*reaction*), lalu dilanjutkan ke level pembelajaran (*learning*), level perilaku (*behavior*) dan terakhir level hasil (*result*) (Kirkpatrick & Kirkpatrick, 2006, pp. 21-25).

Four levels of evaluation that progress from minimal to comprehensive: (1) positive reactions to training, (2) achievement of learning objectives, (3) transfer of learning into behavior change, and (4) explicit identification of results. (Prywes, 2012, p. 43). Ilustrasi model Kirkpatrick dapat dilihat pada Gambar 1.



Sumber: (Prywes, 2012, p. 43)

Gambar 1. Model Evaluasi Program Model Kirkpatrick

Pengumpulan data yang digunakan adalah angket, observasi, wawancara dan studi dokumentasi pada semua level baik Level 1 (reaksi), Level 2 (pembelajaran), Level 3 (perilaku) maupun Level 4 (dampak). Sebelum digunakan, instrumen-instrumen tersebut dilakukan validasi pakar dan konten (isi) yang melibatkan 5 Ahli dalam bi-

dang instrumen, konten widyaiswara dari berbagai instansi.

Komponen yang diukur instrumen pada level 1 (reaksi) terdiri dari kurikulum diklat, mata diklat dan tujuan pembelajaran, kompetensi widyaiswara, sarana dan prasarana (sarpras) dan penunjang pembelajaran serta fasilitas asrama. Komponen kurikulum diklat meliputi kriteria lama belajar, waktu belajar, penjadwalan, mata diklat dan jumlah peserta. Komponen Mata Diklat meliputi kriteria kesesuaian materi dan kesesuaian dengan kompetensi yang dibutuhkan. Komponen Kompetensi Widyaiswara meliputi kepangkatan dan gelar kesarjanaan, metode pembelajaran, media pembelajaran dan evaluasi. Komponen sarpras meliputi kriteria modul, sarana pembelajaran, sarana penunjang dan konsumsi. Komponen Asrama meliputi fasilitas asrama dan sarana penunjang. Skala atau standar pengukuran level 1 (reaksi) didasarkan pada Program Diklat tahun 2013 yang dikeluarkan oleh Pusbangtendik.

Pada instrumen level 2 (Pembelajaran) terdiri dari komponen sikap dan keterampilan. Adapun komponen sikap meliputi kriteria pembetulan sikap ilmiah dengan rerata skor minimal 3,67. Sedangkan komponen keterampilan meliputi kriteria menyusun laporan KTI dengan skor minimal 70.

Pada instrumen level 3 (Perilaku) terdiri dari komponen sikap ilmiah dan perilaku kerja. Komponen sikap ilmiah meliputi kriteria berperan aktif dalam 1 kegiatan ilmiah di unit kerja dalam kurun waktu 6 bulan. Kriteria pada perilaku kerja meliputi penerapan pengetahuan KTI 1 kali dalam kurun waktu 6 bulan dan bekerja sama dalam memprakarsai 1 kegiatan ilmiah di unit kerja dalam kurun waktu 6 bulan.

Pada instrumen level 4 (Dampak) terdiri dari komponen keterampilan menyusun KTI dan jumlah publikasi. Kriteria pada komponen keterampilan menyusun KTI meliputi tersusunnya KTI dan membimbing *stakeholders* dalam menyusun KTI dalam kurun waktu 6 bulan. Kriteria pada jumlah publikasi meliputi publikasi ilmiah minimal berskala lokal dalam kurun waktu 6 bulan.

Skala pengukuran atau standar KTI widyaiswara didasarkan pada Peraturan No. 9 tahun 2008 tentang Pedoman Penyusunan Karya Tulis Ilmiah bagi Widyaiswara.

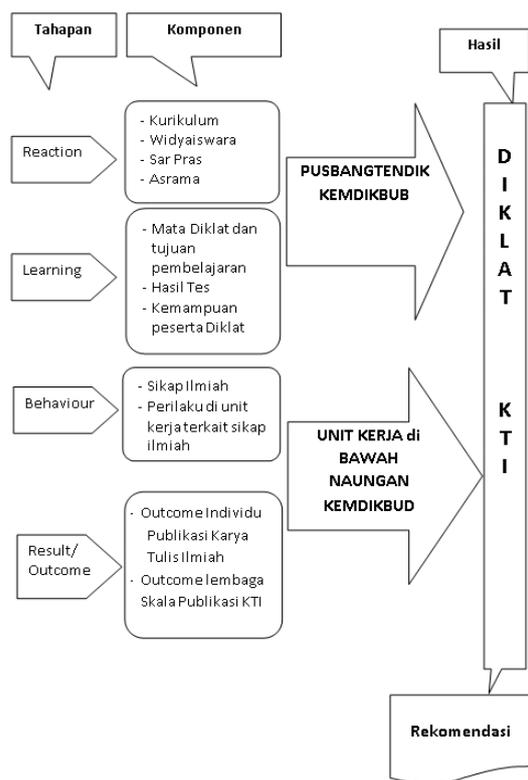
Triangulasi dan pengecekan keabsahan data dilakukan pengecekan silang (*cross check*) kepada widyaiswara pengajar, panitia, teman sejawat dan atasan. Pihak-pihak yang diikutsertakan dalam pengumpulan data meliputi kepala bidang evaluasi dan program pusbangtendik, widyaiswara pengajar, widyaiswara peserta, panitia, rekan sejawat dan pimpinan dari peserta diklat. Hal ini ditujukan agar data-data yang didapat benar-benar sah sesuai pemeriksaan keabsahan data dalam penelitian yang mencakup 4 tahap, yaitu kredibilitas, keterangan, kebergantungan dan kepastian.

Teknik analisis data yang digunakan dalam penelitian ini menggunakan teknik analisis secara umum adalah kualitatif deskriptif dengan mendasarkan keputusan pada kriteria yang telah disusun. Jenis teknik analisis data yang digunakan merujuk pada Miles & Huberman dengan alur reduksi data, sajian data dan penarikan kesimpulan yang disertai verifikasi serta interpretasi data. (Miles & Huberman, 1992, pp. 16–20). Adapun penggunaan statistik deskriptif seperti nilai rata-rata, nilai maksimum dan minimum pada penelitian ini dimaksudkan agar data kuantitatif yang didapatkan lebih mudah dikomparasikan dengan kriteria-kriteria evaluasi dan dikategorikan menjadi Baik, Cukup atau Kurang. Gambaran umum desain penelitiannya disajikan pada Gambar 2.

Hasil Penelitian dan Pembahasan

Hasil Penelitian

Evaluasi atau pengambilan keputusan terhadap empat langkah Kirkpatrick terdiri dari reaksi (*reaction*), hasil belajar (*learning*), perilaku kerja (*behavior*), dan dampak (*result*), dilakukan dengan membandingkan kriteria dan indikator dengan data-data di lapangan. Adapun kriteria tersebut disusun berdasarkan kriteria kuantitatif dan kualitatif dengan pertimbangan (Arikunto & Jabar, 2014, pp. 34–37).



Gambar 2. Desain Evaluasi Program Diklat KTI

Pertimbangan dan pengambilan keputusan tidak selalu bersifat kuantitatif tergantung pada rujukan kriteria yang akan diukur. Adapun skor kuantitatif umumnya didapatkan dari angket, observasi dan studi dokumentasi, sedangkan data kualitatif didapatkan dari observasi, wawancara dan studi dokumentasi.

Pengkategorian baik, jika rentang pengukuran di atas 66%, cukup untuk rentang 56-65% dan kurang jika didapatkan pemenuhan kriteria di bawah 45% baik secara kuantitatif maupun secara kualitatif pada pengukuran nilai, skor & ketercapaian indikator dari sebuah kriteria. Pada beberapa kriteria memang telah tercantum batas minimum kecukupan dari kriteria agar mempermudah pengukuran dan penilaian. Hal tersebut juga untuk mengakomodir ketetapan dan peraturan yang ditetapkan oleh LAN dan Pusbangtendik. Untuk lebih jelasnya dapat dilihat pada tabel kriteria masing-masing level pada tabel 2, 5, 6 dan 7.

Peserta diklat pada periode ini berjumlah 32 orang, yang terdiri dari widyaiswara yang ditugaskan di LPMP dari berbagai provinsi, P4TK dan Pusbangtendik yang keseluruhan berjumlah 23 unit kerja. Kegiatan diawali dengan acara pembukaan dan penjelasan program. Pemberian materi diklat dilakukan di gedung budaya yang difasilitasi oleh Widyaiswara pengajar sesuai dengan jadwal yang telah ditentukan. Setelah itu barulah Diklat KTI secara resmi dimulai.

Evaluasi Reaksi

The first level is the Reaction level in which the reactions of the trainees are understood to mean the way in which they perceive and subjectively evaluate the relevance and quality of the training. According to Kirkpatrick, every program should at least be evaluated at this level to provide for the improvement of a training program. At this level, Evaluation measures the satisfaction of the people who followed the training. (Farjad, 2012, p. 2838). Pada evaluasi Perilaku kerja program KTI kriteria evaluasi dapat dilihat pada Tabel 2.

Berdasarkan reaksi atau tanggapan peserta atas pelaksanaan program Diklat KTI yang berkaitan dengan komponen kurikulum dari mulai waktu, jadwal, mata diklat sampai jumlah peserta pada umumnya dinilai baik.

Diklat KTI terdiri dari 7 hari pelaksanaan diklat, 55 jam pelajaran diklat, 45 menit per jam pelajaran dan 10 jam pelajaran per hari telah sesuai dengan kriteria serta dapat dibuktikan dengan dokumen terkait. Begitu pun dalam hal penjadwalan yang diatur oleh panitia.

Hasil studi dokumentasi didapatkan data bahwa jadwal yang dikeluarkan oleh Pusbangtendik, acara dimulai hari Senin tanggal 9 Maret 2015 dengan diawali oleh *check in* peserta dan pembukaannya dilakukan pada tanggal 10 Maret 2015 dibuka oleh Kepala Bidang Pengembangan Tenaga pimpinan dan Pegawai. Diklat berlangsung sampai dengan hari minggu 15 Maret 2015, dengan penutupan di hari sabtu 14 Maret 2015.

Tabel 2. Kriteria Evaluasi Reaksi Program Diklat

Tahapan	Komponen	Kriteria
Reaksi (Reaction)	Kurikulum Diklat	<ol style="list-style-type: none"> 1. Durasi / lama belajar efektif (45 menit/ mata diklat, 10 JP/hari) 2. Waktu belajar efektif 55 jam pelajaran (7 hari). 3. Kegiatan diklat sesuai dengan jadwal yang diberikan kepada peserta (penjadwalan). 4. Mata diklat sesuai dengan kurikulum yang dikeluarkan LAN. 5. Peserta maksimum terdiri dari 30 orang peserta per kelas.
	Mata Diklat dan tujuan pembelajaran	<ol style="list-style-type: none"> 1. Materi diklat yang diajarkan 100% sesuai dengan kurikulum diklat yang dikeluarkan LAN. 2. Mata diklat memberikan kompetensi yang dibutuhkan sesuai dengan profil peneliti Widyaiswara yang ditentukan oleh LAN sesuai dengan peraturan PERKALAN dan PERKABKN nomor 1 tahun 2015 dan nomor 8 tahun 2015 tentang jabatan fungsional Widyaiswara dan angka kreditnya.
	Kompetensi Widyaiswara (pengajar)	<ol style="list-style-type: none"> 1. Kompetensi widyaiswara sesuai dengan syarat-syarat yang ditetapkan oleh LAN (minimal S1, dan jabatan minimal widyaiswara madya). 2. Metode pembelajaran yang dikembangkan Widyaiswara berlangsung sesuai pedoman penyelenggaraan diklat dengan syarat minimal adanya sesi pembuka, diskusi, tanya jawab dan penarikan kesimpulan mata diklat. 3. Widyaiswara menggunakan minimal 1 media pembelajaran 4. Melakukan evaluasi hasil diklat dengan materi evaluasi sesuai jadwal

Berdasarkan keterangan dari panitia tidak ada pengurangan waktu jam pelajaran bagi peserta diklat. Sehingga peserta juga tidak dirugikan dengan jadwal yang terkesan agak senggang dihari pertama dan terakhir. Keterangan dari panitia tersebut, dapat pula dibuktikan dengan durasi 45 menit per jam pelajaran yang tertera di dokumen penjadwalan diklat. Adapun pemadatan yang dilakukan atas dasar mempertimbangkan efektivitas waktu dan biaya.

Uraian tersebut diperkuat oleh hasil angket yang diberikan kepada peserta. Hasil rata-rata skor angket adalah 4,25 (dengan skala 1-5) yang menyatakan bahwa semua peserta setuju bahwa pelaksanaan Diklat KTI sesuai dengan jadwal. Artinya, kriteria ini dapat dikatakan baik.

Berdasarkan penelusuran modul maupun di dalam berkas jadwal pembelajaran didapatkan hasil 100% kurikulum yang digunakan adalah kurikulum yang berasal dari LAN. Begitu pun dalam berkas Surat Tanda Tamat Pelatihan (STTP) keseluruhan mata diklat yang dinilai pun merujuk pada kurikulum tersebut.

Pada kriteria jumlah peserta melebihi kapasitas yang seharusnya 30 peserta diikuti oleh 32 peserta. Perlu adanya ketegasan dari pihak panitia. Adanya penambahan peserta ini dikarenakan permohonan dari beberapa

unit kerja di bawah naungan Kemdikbud di luar rencana yang telah ditetapkan.

Kondisi tersebut berimbas terhadap kenyamanan belajar peserta lain yang datang berdasarkan undangan dan rekomendasi unit kerjanya. Imbas yang terjadi kelas yang menjadi kurang representatif untuk diklat, kecukupan sarana pembelajaran seperti *printer*, koneksi internet yang semakin padat dan perubahan format-format administratif seperti kehadiran, penilaian maupun blangko sertifikat yang sedianya telah dibuat panitia. Di samping itu, pembagian kelompok menjadi lebih banyak, sehingga dibutuhkan pemadatan waktu bagi kelompok lain yang akan melaksanakan presentasi kelompok.

Pada kriteria tim pengajar yang terdiri dari 2 widyaiswara per mata diklat direspon baik melalui kuisisioner oleh peserta, panitia maupun koordinator widyaiswara. Berlainan dengan data administratif tersebut, data yang diberikan responden melalui wawancara mendalam ditemukan tidak-nyamanan dalam pembelajaran dikarenakan kesenjangan kompetensi dari widyaiswara pengajar yang dirasakan oleh peserta, namun sejauh penilaian secara keseluruhan dapat dikatakan memenuhi kriteria administratif.

Studi dokumentasi yang dilakukan pada evaluasi pembelajaran menunjukkan bahwa seluruh (100%) materi yang diajarkan meru-

pakannya materi yang tercantum dalam kurikulum. Hal ini berarti panitia maupun widyaiswara telah menjalankan tugasnya secara baik.

Pada kriteria kesesuaian mata diklat dan kompetensi yang dibutuhkan sesuai dengan profil widyaiswara pengumpulan data dilakukan dengan angket dan wawancara baik kepada peserta maupun panitia. Wawancara terhadap peserta mengatakan: *“Sangat menunjang itu (Diklat KTI), bagi kami merupakan kebutuhan kami, kita sebagai WI kan menyusun angka kredit, sebaiknya di-link-kan ke sana, tinggal kita menindaklanjuti saja bagaimana menyusun pendahuluan dll yang sesuai tupoksi kita”*

Hasil angket terhadap peserta setelah diklat berlangsung, didapatkan skor sebesar 4,109 (skala 1-5). Hal ini menandakan bahwa materi diklat sesuai dengan kompetensi yang dibutuhkan sebagai seorang widyaiswara.

Kompetensi widyaiswara terdiri atas 4 indikator beserta kriteria evaluasinya. Keempat indikator tersebut terdiri atas syarat widyaiswara, metode pembelajaran yang dikembangkan, media pembelajaran yang digunakan dan evaluasi Diklat KTI. Berdasarkan penilaian yang dilakukan panitia skor widyaiswara dalam penggunaan metode pembelajaran adalah 87,8 (lihat Tabel 3) dari rentang 0-100. Sedangkan berdasarkan angket yang disebarkan kepada peserta, dalam penggunaan metode pembelajaran widyaiswara mendapatkan skor rata-rata 4,047.

Uraian tersebut di atas, dapat diartikan bahwa pengembangan metode pembel-

ajaran dalam Diklat KTI telah memenuhi standar yang dipersyaratkan dan mendapat respon yang baik dari peserta.

Tanggapan peserta terhadap media pembelajaran yang digunakan oleh widyaiswara pada Diklat KTI, umumnya memberikan respon yang baik. Hal ini terlihat dari jawaban angket peserta yang diberikan sesaat setelah diklat berlangsung mendapatkan skor rata-rata 4,125 (skala 1-5) yang berarti penggunaan media pembelajaran oleh widyaiswara tergolong baik. Hal ini didukung oleh data Pusbangtendik yang mengatakan kemampuan widyaiswara dalam menyajikan mendapatkan skor 86,9 dan 86,9 (lihat Tabel 3).

Melakukan Evaluasi pembelajaran Diklat KTI merupakan wewenang dari LAN sebagai regulator Diklat. Oleh karena itu, dalam indikator melakukan evaluasi pembelajaran digunakan data hasil belajar yang sudah ada (data sekunder) melalui studi dokumentasi yang dikaitkan dengan respon peserta, panitia dan widyaiswara itu sendiri melalui angket.

Berdasarkan studi dokumentasi tentang pelaksanaan Diklat KTI. Evaluasi pembelajaran yang seharusnya dilakukan Sabtu sore tanggal 14 Maret 2015. tepat setelah selesai mata diklat terakhir. Hal ini sesuai dengan jadwal yang telah dirancang oleh panitia sebelumnya.

Tabel 3. Rekapitulasi penilaian Widyaiswara Pusbangtendik pada Diklat KTI.

No	Komponen penilaian	Nilai
1	Penguasaan materi	87,5
2	Sistematika penyajian	86,9
3	Kemampuan menyajikan	86,9
4	Relevansi materi dengan tujuan pembelajaran	87,7
5	Penggunaan metode belajar dan sarana diklat	87,8
6	Penggunaan Bahasan	88,3
7	Nada dan suara	86,9
8	Cara menjawab pertanyaan peserta	86,6
9	Gaya sikap dan perilaku	86,7
10	Pemberian motivasi kepada peserta	88,1
11	Kualitas bahan diklat	86,9
12	Kerapihan berpakaian	87,7
13	Disiplin kehadiran	88
14	Kerja sama antar Widyaiswara	89,2
Nilai rata-rata Kompetensi Widyaiswara Diklat KTI		87,36

Sumber: (Pusbangtendik, 2015)

Kecukupan sarana dan pra-sarana pembelajaran dinilai memenuhi standar. Kelayakan pakai, kondisi dan jumlah dari sarana prasarana semuanya menunjang pembelajaran. Di ruang belajar dilengkapi LCD, perangkat alat tulis, papan *White board*, printer, *Wi Fi* dan *AC*. Terlepas dari rasio printer dan koneksi internet melalui *Wi Fi* yang masih dinilai kurang dalam hal kecepatan dan kapasitasnya. Namun demikian, masih dapat dikatakan wajar. Berdasarkan uraian tersebut di atas, maka indikator evaluasi pembelajaran pada dasarnya dilaksanakan sesuai jadwal. Artinya, pembelajaran mata diklat telah selesai kemudian baru diadakan evaluasi pembelajaran diklat.

Adapun hal lain yang perlu dibenahi adalah ketersediaan air minum berupa galon air mineral di sekitar area asrama, sehingga peserta tidak perlu lagi mencari atau membeli air mineral dan dibawa ke asrama. Terlebih jarak antara asrama dan dapur melebihi 500 meter yang terbilang jauh untuk sekedar meminta air minum di malam hari. Berdasarkan penilaian yang dilakukan oleh Pusbangtendik sendiri atas proses pembelajaran yang dilakukan oleh widyaiswara menunjukkan didapatkan hasil penilaian yang

terbilang sudah baik dengan skor 87,36 (rentang 0-100). Adapun rinciannya disajikan pada Tabel 4.

Untuk indikator asrama berupa kamar yang representatif, kamar mandi hingga peralatan tidur yang diberikan kepada peserta Diklat KTI telah memenuhi syarat kriteria 1:1. Terlebih dengan adanya bagian informasi yang sigap memberikan informasi dan petugas kebersihan yang menjalankan tugasnya secara periodik merupakan hal yang membuat peserta puas dalam pelayanan pada aspek reaksi ini.

Di samping itu, atas dasar penilaian angket yang dilakukan oleh Pusbangtendik sendiri didapatkan penilaian yang terbilang sudah baik dengan skor 85,93 (rentang 0-100). Adapun rinciannya disajikan pada Tabel 4.

Evaluasi Pembelajaran

Learning can be described as the extent to which the attitudes of the participants change, their knowledge increases or their skills are broadened as a consequence of the training. (Farjad, 2012, p. 2838) Kriteria Evaluasi pembelajaran program KTI dapat dilihat pada Tabel 5.

Tabel 4. Rekapitulasi Penilaian Sarana dan prasarana Diklat KTI oleh Pusbangtendik

No	Komponen penilaian	Nilai
1	Kelas/ Gedung Budaya	87,0
2	Asrama	82,7
3	Kantin Paramita	87,9
4	Fasilitas olah raga	86,2
5	Fasilitas Unit kesehatan	82,6
6	Fasilitas Masjid/Musholla	84,6
7	Fasilitas Hiburan/Refreshing dalam kampus	86,5
8	Fasilitas penerangan kampus	87,6
9	Kebersihan lingkungan kampus	88,3
Nilai rata-rata Layanan Sarana dan prasarana Pusbangtendik		85,93

Sumber: (Pusbangtendik, 2015)

Tabel 5. Kriteria Evaluasi Pembelajaran Diklat KTI

Tahapan	Komponen	Kriteria
Hasil Belajar (<i>Learning</i>)	Hasil Pembelajaran dalam ranah sikap ilmiah	Pembentukan sikap ilmiah yang diukur melalui skala dengan syarat minimal cukup yang dinilai oleh tim penilai dari panitia (minimal rata-rata skor 21 atau $\geq 3,67$).
	Hasil pembelajaran dalam ranah keterampilan menyusun KTI	Mampu menyusun Laporan Karya Tulis Ilmiah sesuai dengan pedoman penyusunan karya tulis ilmiah untuk widyaiswara yang dikeluarkan LAN, dibuktikan dengan penilaian akhir diklat (minimal skor total 70).

Evaluasi pembelajaran Diklat KTI merupakan wewenang dari LAN sebagai regulator diklat, baik jadwal, penilaian akhirnya maupun materi tes yang dibuat. Oleh karena itu, dalam indikator evaluasi pembelajaran digunakan data sekunder melalui studi dokumentasi yang dikaitkan dengan respon peserta, panitia dan widyaiswara itu sendiri melalui angket.

Widyaiswara pengajar dari LAN kali ini tidak memberikan tes awal (*pre test*) pada peserta. *Pre-test* yang ada, diberikan oleh Pusbangtendik sebagai operator untuk kepentingan pemetaan peserta. Pengukuran dan penilaian diklat dititikberatkan pada pencapaian peserta untuk memenuhi syarat Sikap ilmiah, kemampuan menyusun KTI dan keterampilan menyajikan dalam bentuk tulisan serta presentasi.

Rata-rata dari 32 peserta pada *pre test* didapatkan skor 72,5, sedangkan hasil *post test* didapatkan skor rata-rata sebesar 82,95. Skor-skor tersebut menandakan bahwa terdapat peningkatan yang cukup tajam, sehingga pada level pembelajaran dapat dikategorikan berhasil dengan baik.

Pada aspek mata diklat dan tujuan pembelajaran, panitia mengikuti kurikulum yang telah ditetapkan oleh regulator dari LAN. Mata diklat yang diturunkan dari kurikulum dan kompetensi yang dibutuhkan oleh peserta yang merupakan widyaiswara di Kemdikbud, secara umum telah disusun sedemikian rupa menjadi rangkaian proses pembelajaran yang telah terstandarkan oleh panitia.

Aspek hasil belajar yang terdiri dari penilaian sikap ilmiah yang mencakup 30% dari seluruh penilaian, pengetahuan dan keterampilan menyusun KTI yang mencakup 70% dari seluruh penilaian. Pada penilaian sikap dilakukan seiring berjalannya proses pembelajaran oleh widyaiswara pendamping. Sedangkan indikator pengetahuan dan keterampilan dilakukan oleh widyaiswara pengajar.

Hasil evaluasi pembelajaran Diklat yang telah dirangkum dari berbagai penilaian dan kemudian dibandingkan dengan kriteria minimum yang telah ditentukan oleh LAN sebagai regulator. Nilai minimal sebesar 70 adalah kriteria minimal kelulusan dari Diklat KTI ini. Hasil diklat menunjukkan bahwa nilai paling kecil adalah 88,81.

Berdasarkan hasil dan kriteria yang telah ditetapkan dapat ditarik kesimpulan 100% atau seluruh peserta dinyatakan lulus dengan berbagai predikat. Artinya, tujuan pembelajaran yang telah ditetapkan oleh LAN melalui kurikulum, modul dan mata diklat dapat dicapai oleh seluruh peserta.

Evaluasi Perilaku Kerja

Pada level *behavior* ini berusaha untuk mengevaluasi seberapa banyak “*transfer of knowledge, skills, and attitudes*” terjadi sebagai akibat dari program diklat yang diikutinya. (Wijaya & Sumarno, 2017, p. 132). Adapun evaluasi perilaku kerja program KTI kriteria evaluasi dapat dilihat pada Tabel 6.

Tabel 6. Kriteria Evaluasi Perilaku Kerja Diklat KTI

Tahapan	Komponen	Kriteria
Perilaku Kerja (<i>behavior</i>)	Sikap ilmiah	Tumbuhnya Sikap ilmiah dengan ditandai adanya peran serta aktif dalam kegiatan ilmiah di dalam unit kerja minimal 1 kali dalam 6 bulan.
	Perilaku kerja	<ol style="list-style-type: none"> 1. Penerapan pengetahuan dan aplikasi KTI di unit kerja minimal menyusun KTI non penelitian (membimbing penyusunan KTI, menyusun bahan ajar, modul, dll) 1 kali dalam 6 bulan. 2. Bekerja sama dan memprakarsai kegiatan ilmiah sesama widyaiswara di unit kerja minimal 1 kegiatan dalam 6 bulan.

Hambatan dalam menerapkan ilmu yang didapat dalam Diklat KTI cukup beragam, sehingga pendekatan yang dilakukan untuk menginisiasi pada alumnus juga berbeda. Setelah 6 bulan berlalu, efektivitas Diklat KTI terlihat mulai menghasilkan perilaku kerja yang cukup baik. Perilaku yang demikian dikarenakan umumnya alumnus langsung melakukan perbaikan atas karya ilmiah mereka yang ditulis pada waktu Diklat. Artinya, tambahan pengetahuan dan keterampilan yang didapat langsung dipraktikkan dalam skala individual dan terbatas.

Hasil observasi dan wawancara kepada rekan sejawat dan atasan langsung, mengisyaratkan bahwa pentingnya iklim ilmiah dan dukungan dari lingkungan kerja, ini terbukti di beberapa unit kerja yang begitu responsif terhadap karya tulis ilmiah dari alumnus atau widyaiswara, tapi di unit kerja lain terlihat kurang sekali responnya terhadap karya tulis ilmiah dari alumnus atau widyaiswara.

Kedua kondisi tersebut menegaskan penerapannya hasil pembelajaran di unit kerja masih terkendala faktor unit kerja. Artinya, dibutuhkan dukungan dari unit kerja dan juga pimpinan di unit kerja masing-masing ketika alumnus kembali ke unit kerjanya. Dukungan dapat diberikan dengan diberikan dorongan melakukan penelitian bagi para alumnus, memfasilitasi alumnus dengan jurnal, majalah dan menulis bahan ajar atau modul melalui mekanisme hibah.

Evaluasi Dampak

The measure key performance of trainee and what is change in the organization after training like staff turnover rate, employee retention, quality ratings and other types of quantifiable aspects of the performance of the organization. (Srivastava & Agarwal, 2012, p. 9). Adapun kriteria evalu-

asi dampak Diklat KTI dapat dilihat pada Tabel 7.

Data penelitian menunjukkan 12 orang alumnus Diklat KTI yang sudah berhasil menyusun karya tulis ilmiah disertai pengakuan unit kerja atau tercatat di unit kerjanya masing-masing. Dari 12 orang tersebut, hanya 8 alumnus yang berhasil mempublikasikannya.

Studi dokumentasi yang dilakukan setelah 6 bulan pelaksanaan Diklat KTI menemukan hal yang unik. Terdapat 5 orang alumnus yang tergolong memiliki pencapaian hasil diklat terendah justru telah memiliki karya ilmiah, dari 5 orang tersebut, 3 diantaranya melakukan publikasi ilmiah ke jurnal. Sebaliknya, sepuluh peserta dengan pencapaian hasil belajar tinggi, hanya 4 orang yang berhasil menyusun karya tulis ilmiah. Keempat orang tersebut, 2 diantaranya melakukan publikasi ilmiah ke dalam jurnal. Tentunya ini cukup mengherankan dan bertentangan dengan hasil kajian Amaddin, Fitriyah & Irawan (2015, p. 158) yang menyatakan bahwa peserta diklat yang memiliki prestasi baik pada pembelajaran, maka akan baik pula pada tingkat kinerjanya.

Apabila ditinjau dari sisi *outcome* lembaga dari 23 unit kerja yang mengirimkan peserta baru 8 unit kerja yang mempublikasikan KTI-nya ke dalam bentuk jurnal, modul maupun laman dari unit kerjanya masing-masing. Dengan demikian, dapat disimpulkan bahwa pada komponen evaluasi dampak Diklat KTI tergolong kurang memberikan dampak baik secara individu maupun kelembagaan. Oleh karena itu, diperlukan upaya yang lebih keras lagi dalam mendorong unit kerja untuk memfasilitasi dan menerbitkan jurnal, modul maupun media publikasi lainnya.

Tabel 7. Kriteria Evaluasi Dampak Diklat KTI

Tahapan	Komponen	Kriteria
Dampak (Result)	Keterampilan menyusun karya tulis ilmiah	1. Tersusunnya Karya Tulis Ilmiah yang tercatat di unit kerja, minimal 1 buah setelah diklat. 2. Membimbing guru atau stakeholder unit kerja dalam menyusun karya tulis ilmiah minimal 2 orang guru per Widyaiswara.
	Jumlah Publikasi karya tulis ilmiah	Publikasi ilmiah berskala lokal, nasional atau internasional dalam kurun waktu 6 bulan setelah diklat minimal 1 buah publikasi.

Pembahasan

Kondisi yang telah diungkapkan di atas, pada dasarnya bukan merupakan hal yang aneh. Amaddin, Fitriyah dan Irawan menemukan hal yang senada dalam kajiannya, terbatasnya staf pengajar/widyaiswara yang profesional, atau memenuhi kualifikasi yang dibutuhkan, mengakibatkan kurang efektifnya proses belajar mengajar (Amaddin, Fitriyah, & Irawan, 2015, p. 158)

Lebih lanjut dikemukakan Hidayat & Sa'ud (2015, p. 30) bahwa kompetensi widyaiswara terkait dengan proses pembelajaran pada kegiatan pendidikan dan pelatihan masih terjadi *miss match*, diantaranya terdapat widyaiswara yang mengampu mata diklat yang tidak sesuai dengan latar belakang pendidikannya, tetapi ada juga widyaiswara yang menguasai semua ilmu. Hal ini berdampak pada kualitas pelaksanaan pendidikan dan pelatihan.

Sarana dan prasarana yang berupa alat cetak *printer* dan koneksi internet serta ketersediaan air minum di asrama menjadi suatu hal yang perlu diperhatikan oleh Pusbangtendik terkait pelayanannya dalam sarana dan prasarana. Pelayanan diklat dinilai oleh peserta sudah baik. Pelayanan di ruang kelas yang representatif, ruang fasilitator, kelengkapan diklat seperti modul dan alat tulis hingga pelayanan dalam konsumsi dapat memuaskan peserta diklat.

Tujuan dari pembelajaran pada dasarnya adalah terlaksananya kurikulum dan hasil belajar seluruh peserta di atas standar serta kelulusan 100%. Pada level pembelajaran didapatkan rata-rata skor kelulusan 82,95 pada kriteria kelulusan 70 sehingga dapat disimpulkan bahwa aspek tersebut telah memenuhi regulasi dan persyaratan yang ditetapkan. *Learning criteria are measures of the learning outcomes of training; they are not measures of job performance. They are typically operationalized by using paper-and-pencil and performance tests.* (Arthur et al., 2003, p. 235)

Kemampuan beberapa peserta sudah jauh lebih berkembang dan berhasil meningkatkan sikap ilmiah, pengetahuan maupun keterampilannya. Peningkatan dari skor *pre-test* ke *post-test*, tidak adanya complain

dari peserta terhadap kurikulum, modul dan proses pembelajaran merupakan bukti bahwasanya proses yang dijalankan sesuai kriteria.

Untuk level perilaku kerja dukungan dari unit kerja mejadi faktor penghambat bagi alumnus dalam mengimplementasikan hasil pembelajaran dalam Diklat KTI. Sehingga level perilaku kerja yang ditandai dengan munculnya sikap ilmiah, dan penerapan pengetahuan serta keterampilan dari Diklat KTI dapat dikatakan cukup. Hal ini mendukung pernyataan Arthur bahwa *behavioral criteria are measures of actual on-the-job performance and can be used to identify the effects of training on actual work performance.* (Arthur et al., 2003, p. 158)

Perhatian dan dukungan yang lebih kongkrit dari unit kerja dan Pusbangtendik untuk melakukan monitoring dan evaluasi merupakan tindakan yang ditunggu oleh alumnus. Berdasarkan data tidak semua unit kerja melakukan dukungan penuh terhadap aktualisasi diri dalam melaksanakan penyusunan karya tulis, sehingga butuh dilakukan dukungan secara langsung dari pihak Pusbangtendik.

Kesinambungan model evaluasi program diklat pada level ini mulai terganggu dikarenakan selain kompetensi dari alumnus, ternyata ada faktor lain yang cukup mengganggu, yaitu lingkungan kerja dan kebijakan unit kerja. Beberapa wawancara kepada rekan kerja dan atasan alumnus di unit kerja membuktikan perhatian terhadap karya tulis ilmiah secara masih jauh dari harapan. Padahal selayaknya pendidik bagi guru, dosen maupun widyaiswara memiliki tugas yang hampir sama dalam hal pengembangan profesi, yaitu menulis karya ilmiah.

Uraian tersebut di atas sekaligus mengkonfirmasi penelitian yang dilakukan Anisah et al (2015, p. 9) yang menyatakan bahwa publikasi ilmiah yang dilakukan widyaiswara tingkat nasional (kecuali pulau jawa) hanya mencapai skor 6,3 untuk pulau sumatera, 5,5 untuk kalimantan, 7,2 untuk Bali dan nusa tenggara Barat, 2,9 untuk Maluku dan Papua dari rentang skor 1-10.

Contoh sederhana, terdapat peserta dengan skor terendah pada Diklat KTI ini (77,28) yang memiliki hasil evaluasi perilaku kerja dan publikasi, peserta ini lebih unggul dibanding alumnus lainnya. Ia telah mempublikasikan buku bersama dengan sejawatnya di bawah naungan LPMP Bali dan mempublikasikan penelitiannya melalui Jurnal HEPI (Himpunan Evaluasi Pendidikan Indonesia) cabang Bali.

Senada dengan keterangan di atas, Saad & Mat (2013, p. 20) mengatakan *the measurement is based on the notion that training and human resources development must reflect the organizational culture and strategy. A training program is judge successful only if the training outcome aligned closely with the organization's goals*. Lingkungan kerja dan motivasi memberi pengaruh langsung kepada kinerja pegawai balai Diklat Keagamaan Manado (Paita, Tewal, & Sendow, 2015, p. 83).

Kedua hasil di atas, menjelaskan bahwa LPMP Bali mempunyai suasana yang kondusif dan mendukung dalam iklim penelitian bagi widyaiswara. Hasil ini mengisyaratkan pula, lingkungan kerja yang kondusif berdampak baik atas peningkatan prestasi individu widyaiswara.

Berdasarkan pendalaman yang dilakukan ternyata didapatkan bahwa tiap-tiap Unit kerja LPMP, P4TK maupun Pusbangtendik memiliki perhatian yang berbeda satu dengan yang lainnya terhadap publikasi karya ilmiah, hal ini terbukti hanya beberapa unit kerja yang memiliki jurnal, majalah maupun kolom publikasi artikel ilmiah di laman unit kerjanya. Pada salah satu unit kerja juga didapatkan bahwa tidak ada dukungan maupun pendanaan sama sekali untuk kegiatan penelitian yang dilakukan widyaiswara.

Hal tersebut di atas mengkonfirmasi pernyataan Bates (2004, p. 342) yang mengatakan *There are at least three limitations of Kirkpatrick's model that have implications for the ability of training evaluators to deliver benefits and further the interests of organizational clients. These include the incompleteness of the model, the assumption of causality, and the assumption of increasing importance of information as the levels of outcomes are ascended*.

Badu mengatakan (2012, pp. 102–129) untuk mendapatkan gambaran tentang implementasi Evaluasi Model Kirkpatrick. Keempat level untuk Evaluasi Model Kirkpatrick masing-masing dilakukan penilaian dengan alat penilaian yang berbeda. Penilaian kinerja digunakan untuk menilai hasil kerja atau proyek menggambarkan Evaluasi Model Kirkpatrick.

Dampak atau *outcome* dari diklat dapat dibagi menjadi 2, yaitu *outcome* bagi individu dan lembaga. Kedua *outcome* tersebut, dapat dilihat dari capaian individu tentang hasil karya tulis ilmiah yang masih minim. Setelah 6 bulan Diklat KTI berlangsung dari 32 alumnus baru 14 alumnus yang berhasil mengajukan karya tulis ilmiahnya ke unit kerjanya masing-masing. Ditinjau dari *outcome* lembaga dari 32 alumnus baru 8 alumnus yang mempublikasikan karya tulis ilmiahnya ke jurnal maupun laman dari unit kerjanya masing-masing.

Selain karya tulis ilmiah berbentuk laporan penelitian, widyaiswara juga diwajibkan menulis karya ilmiah dalam bentuk makalah, majalah ilmiah, modul maupun bahan ajar lainnya. Memang dalam waktu 6 bulan dirasa merupakan waktu yang teramat pendek, jika alumnus diminta untuk menyusun itu semua. Namun demikian, dalam waktu 6 bulan pascadiklat, karya tulis yang disusun pada saat diklat dapat ditindaklanjuti menjadi suatu karya yang tercatat di unit kerja sebagai usulan kenaikan pangkat maupun dipublikasikan.

Alie (2015, p. 106) menyatakan bahwa sebagian besar widyaiswara tidak aktif dalam pelaksanaan KTI. Widyaiswara yang aktif menulis, melakukan pelatihan dan memperdalam penulisan karya ilmiah atas inisiatif dari widyaiswara itu sendiri. Sebagian besar widyaiswara menyusun KTI karena hanya menjalankan tugas yang diberikan oleh instansinya untuk mengikuti diklat.

Kajian tersebut memperkuat temuan dalam penelitian evaluasi program KTI bahwasanya widyaiswara umumnya lemah dalam implementasi. Lebih lanjut dinyatakan bahwa salah satu penyebab dari lemahnya widyaiswara adalah dukungan unit kerja

dalam mendorong widyaiswaranya dalam melakukan penelitian, penyusunan dan publikasi KTI.

Simpulan

Secara umum pelaksanaan Diklat KTI di Pusbangtendik Kemdikbud termasuk dalam kategori yang baik pada level 1 (reaksi), level 2 (pembelajaran) dan level 3 (perilaku). Untuk level 4 (dampak) masih dalam kategori kurang dan diperlukan upaya yang keras dari unit kerja dan Pusbangtendik agar widyaiswara meningkatkan jumlah KTI dan publikasinya.

Dukungan terhadap almunus dengan mendorong upaya permbenahan kebijakan dalam hal budaya meneliti, menyusun KTI dan pembiayaan beserta publikasinya di unit kerja masing-masing merupakan suatu hal yang penting. Oleh karena itu, rekomendasi dari evaluasi program Diklat KTI adalah dilanjutkan dengan perbaikan pada dorongan memfasilitasi widyaiswara dalam menulis KTI.

Keberhasilan diklat ternyata tidak dapat berdiri sendiri, peran unit kerja dari masing-masing peserta memiliki peran yang besar dalam keberhasilan diklat. Penggunaan Model Kirkpatrick dalam mengevaluasi Diklat KTI perlu disempurnakan sehingga aspek *outcome* dan dukungan kepada peserta diklat dapat dikembalikan ke unit kerjanya masing-masing.

Daftar Pustaka

Alie, M. (2015). Motivasi Widyaiswara dalam penulisan karya tulis ilmiah (studi kasus pada peserta diklat karya tulis ilmiah di LAN). *Irfani*, 11(1), 96–107.

Amaddin, S., Fitriyah, N., & Irawan, B. (2015). Pendidikan dan pelatihan tot dalam meningkatkan kinerja pegawai widyaiswara di badan pendidikan dan pelatihan provinsi kalimantan timur. *Jurnal Administrative Reform*, 3(1), 148–160.

Anisah, Sukmawati, A., & Made, S. I.

(2015). *Pengaruh pelatihan terhadap kompetensi dan kinerja peneliti*. <http://repository.ipb.ac.id/handle/123456789/79084>. Jakarta.

Arikunto, S., & Jabar, C. S. Abdul. (2014). *Evaluasi program pendidikan* (2nd ed.). Jakarta: Bumi Aksara.

Arthur, R. (2015). Learning approach of problem solving for increase learning achievement of the civil engineering evaluation program. *American Journal of Educational Research*, 3(8), 964–967. <https://doi.org/10.12691/education-3-8-3>

Arthur, W., Bennett, W., Edens, P. S., & Bell, S. T. (2003). Effectiveness of training in organizations: A meta-analysis of design and evaluation features. *Journal of Applied Psychology*, 88(2), 234–245. <https://doi.org/10.1037/0021-9010.88.2.234>

Badu, Syamsu Q. (2012). Implementasi evaluasi model Kirkpatrick pada perkuliahan masalah nilai awal dan syarat batas. *Jurnal Penelitian Dan Evaluasi Pendidikan*, 16, 102–129.

Bates, R. (2004). A critical analysis of evaluation practice: The Kirkpatrick model and the principle of beneficence. *Evaluation and Program Planning*, 27(3), 341–347. <https://doi.org/10.1016/j.evalprogplan.2004.04.011>

Farjad, S. (2012). The evaluation effectiveness of training courses in university by Kirkpatrick model (case study: Islamshahr University). *Procedia - Social and Behavioral Sciences*, 46, 2837–2841. <https://doi.org/10.1016/j.sbspro.2012.05.573>

Hidayat, A. I., & Sa'ud, U. S. (2015). Model pendidikan dan pelatihan berbasis kompetesni bagi widyaiswara muda. *Jurnal Administrasi Pendidikan*, xxii(2), 23–38.

- Kirkpatrick, D. L., & Kirkpatrick, J. D. (2006). *Evaluating training programs*. San Francisco: Berrett-Koehler Publisher, Inc.
- Miles, M. B., & Huberman, A. M. (1992). *Analisis data kualitatif*. Jakarta: UI Press.
- Paita, S., Tewel, B., & Sendow, G. M. (2015). Pengaruh kompensasi dan lingkungan kerja terhadap kinerja pegawai melalui motivasi kerja pada Balai Pendidikan dan Pelatihan Keagamaan Manado. *Jurnal EMBA*, 3(3), 683–694.
- Prywes, Y. (2012). Cognitive, behavioral, and affective learning outcomes of a coaching program. *The International Journal of Mentoring and Coaching*, X(1), 41–55.
<https://doi.org/10.1177/0278364914543793>
- Pusbangtendik. (2015). *Laporan penyelenggaraan diklat kewidyaiswaraan penyusunan karya tulis Ilmiah*. Depok.
- Saad, M. A., & Mat, N. B. (2013). Evaluation of Effectiveness of Training and Development: The Kirkpatrick Model. *Asian Journal of Business and Management Sciences*, 2(11), 14–24.
- Srivastava, E., & Agarwal, N. (2012). Evaluation of training Program. *International Journal of Applied Research & Studies*, I(III), 1–13.
- Wijaya, A., & Sumarno. (2017). Evaluasi dampak pendidikan dan pelatihan pengembangan keprofesian berkelanjutan guru matematika di PPPPTK matematika Yogyakarta. *Penelitian Dan Evaluasi Pendidikan*, 21(2), 127–141.
<https://doi.org/10.21831/pep.v21i2.10113>

KEEFEKTIFAN PERANGKAT PEMBELAJARAN BERBASIS *INQUIRY LESSON* UNTUK MENINGKATKAN KETERAMPILAN BERPIKIR KRITIS SISWA

Susilowati^{1*}, Sajidan¹, Murni Ramli¹

¹Universitas Sebelas Maret

¹Jln. Ir. Sutami No.36A, Jebres, Kota Surakarta, Jawa Tengah 57126, Indonesia

* Corresponding Author. Email: susilotwin@yahoo.co.id

Abstrak

Penelitian ini bertujuan untuk mengetahui efektivitas Perangkat Pembelajaran Biologi untuk meningkatkan keterampilan berpikir kritis siswa. Tempat dan sampel penelitian dipilih secara *purposive random sampling*. Perangkat pembelajaran yang dikembangkan dan diteliti meliputi Rencana Pelaksanaan Pembelajaran (RPP), modul dan instrumen penilaian. Perangkat Pembelajaran tersebut dikembangkan berdasarkan indikator berpikir kritis menurut Facione yang terdiri dari enam aspek, yaitu: interpretasi, analisis, evaluasi, kesimpulan, penjelasan dan pengaturan diri. Instrumen yang dikembangkan, yaitu RPP, modul dan instrumen penilaian berupa tes telah diuji validitasnya dengan menggunakan formula Aiken V dengan indeks validitas secara berturut-turut sebesar 0,93, 0,92, dan 0,90. Adapun estimasi reliabilitas instrumen tersebut adalah 0,66, 0,69 dan 0,74. Hasil Uji efektivitas perangkat pembelajaran menggunakan *T-test* dan diperoleh *p-value* 0,00 ($p < 0,05$) dengan *score gap* antara *pre test* dan *post test* sebesar 31,93%, sehingga dapat disimpulkan bahwa perangkat pembelajaran guru kelas XI pada materi sistem ekskresi manusia efektif untuk meningkatkan keterampilan berpikir kritis siswa.

Kata kunci: *efektivitas, perangkat pembelajaran, kemampuan berpikir kritis*

THE EFFECTIVENESS OF INQUIRY BASED LEARNING TOOLS FOR IMPROVING STUDENTS CRITICAL THINKING SKILLS

Abstract

The research aims to determine the effectiveness of Biology learning device to improve students' critical thinking skills. The research sites and samples were chosen by purposive random sampling. The learning device developed and researched included the lesson plan (RPP), module and assessment instrument. The learning device was developed based on Facione's critical thinking indicator consisting of six aspects, including: interpretation, analysis, evaluation, conclusion, explanation, and self-regulation. The developed instruments, namely the lesson plan, module, and assessment instrument in the form of test had been tested for validity by using Aiken V formula with the validity index of 0.93, 0.92, and 0.90 respectively. The instrument reliability estimation is 0.66, 0.69. and 0.74. The effectiveness of the learning device was tested using T-test and it obtains the p-value of 0.00 ($p < 0.05$) with score gap between pre test and post test at 31.93%, so it can be concluded that the class XI teacher's learning tools on human excretion system material is effective to improve the students' critical thinking skill.

Keywords: *effectiveness, lesson plan, critical thinking skills*

Permalink/DOI: <http://dx.doi.org/10.21831/pep.v22i1.17836>

Pendahuluan

Keterampilan berpikir kritis merupakan salah satu kecakapan berpikir abad 21 yang perlu ditekankan pada bidang pendidikan (The Partnership for 21st Century Skills, 2015, p. 3). Keterampilan berpikir kritis dianggap sebagai pondasi keterampilan lain yang dibutuhkan diantaranya keterampilan komunikasi, keterampilan berkolaborasi, kesadaran global (*global awareness*), keterampilan teknologi, kecakapan dalam hidup dan karir, serta kemampuan belajar serta inovasi. Kemampuan berpikir kritis berguna untuk menstimulasi kemampuan lainnya, seperti berpikir logis, kreatif, memecahkan masalah, berpikir kritis, penguasaan teknologi serta kemampuan beradaptasi terhadap berbagai perubahan dan perkembangan zaman. Huitt (Friedel, Irani, Rudd, Gallo, Eckhardt, & Ricketts, 2008, p.2) menyatakan bahwa berpikir kritis merupakan alat yang penting untuk meraih keberhasilan di abad 21. Zikovic (2016, p. 102) juga menyatakan bahwa model berpikir kritis merupakan atribut penting untuk sukses di abad 21.

Sternberg (1986, p. 6) menyatakan berpikir kritis adalah suatu proses mental, strategi dan representasi yang dilakukan oleh seseorang dalam memecahkan, menyusun keputusan serta dalam mempelajari konsep baru. Norris & Ennis (1986, p. xvii) mendefinisikan berpikir kritis meliputi cara berpikir yang masuk akal (*reasonable*) dan reflektif yang berpusat pada keputusan tentang apa yang selanjutnya dipercaya ataupun dilakukan. Keterampilan berpikir kritis dapat diberdayakan melalui berbagai bidang diantaranya adalah bidang pendidikan. Didalam kegiatan pembelajaran pengembangan berpikir kritis dapat dilatihkan untuk mengembangkan keterampilan berpikir sejak dini.

Pendidikan di Indonesia mengamanatkan bahwa kemampuan berpikir kritis hendaknya dilatihkan pada peserta didik sejak dini. Dalam Kurikulum 2013 menyebutkan bahwa berpikir kritis merupakan salah satu tujuan pendidikan dan menjadi tuntutan kompetensi masa depan yang harus dikembangkan pada peserta didik (Kementerian

Pendidikan dan Kebudayaan, 2016, p. 10). Meskipun keterampilan berpikir kritis merupakan salah satu keterampilan masa depan yang disarankan oleh pendidikan di Indonesia, namun faktanya keterampilan berpikir kritis siswa di Indonesia masih tergolong rendah.

Hasil survei *The Trends in International Mathematics and Science Study* (TIMSS) (National Center for Educational Statistics, 2015, p. 1), menunjukkan bahwa kemampuan skor rata-rata sains peserta didik Indonesia adalah 397. Hasil tersebut jauh dibawah rata-rata internasional sebesar 500 dan menduduki peringkat 4 terbawah dari seluruh negara partisipan. Demikian juga hasil penelitian Programme for International Student Assessment (Balitbang, 2013) diperoleh fakta bahwa kemampuan rata-rata siswa Indonesia dalam bidang matematika dan sains berada pada peringkat 64 dari total 65 negara partisipan. Mendukung temuan di atas, hasil penelitian Suratno & Kurniati (2017, p. 2) menunjukkan bahwa ketertarikan siswa dalam hal membuktikan suatu prinsip maupun konsep, melakukan penyelidikan dan penggeneralisasian masih kurang. Siswa juga kurang tertarik dalam menyelesaikan soal yang bersifat non rutin. Sedangkan hasil penelitian Prihatni, Kumaidi, & Mundilarto, (2016, p. 121) memperoleh fakta bahwa siswa lebih menguasai soal yang berbentuk ingatan dan hafalan tanpa memahami konsepnya. Siswa cenderung hafal dengan rumus-rumus dan tahapan-tahapan penyelesaiannya tanpa memahami konsep dan prosedur yang benar. Fakta tersebut menunjukkan bahwa kemampuan peserta didik Indonesia berada pada *low order thinking*.

Rendahnya pencapaian keterampilan berpikir kritis siswa ini menunjukkan adanya kesenjangan harapan dunia pendidikan dengan fakta yang terjadi di lapangan. Rendahnya keterampilan berpikir kritis siswa disebabkan oleh beberapa faktor yang mempengaruhi proses pembelajaran, bisa dari faktor guru, siswa, maupun faktor pendukung pelaksanaan pembelajaran lainnya. Keterampilan berpikir guru berperan dalam mengembangkan keterampilan berpikir kri-

tis siswa. Demirhan & Koklukaya (2013, p. 1551) dalam penelitiannya memperoleh fakta bahwa pemikiran kritis guru sains tergolong dalam kategori sedang dan rendah

Salah satu faktor pendukung keberhasilan pembelajaran adalah bahan ajar. Bahan ajar didefinisikan sebagai segala bentuk bahan yang disusun secara sistematis untuk membantu guru/instruktur dalam melaksanakan kegiatan pembelajaran dan memungkinkan siswa untuk belajar (Kementerian Pendidikan dan Kebudayaan, 2010, p. 26). Bahan ajar dapat berupa bahan cetak, seperti *hand out*, buku, lembar kegiatan siswa (LKS), modul, brosur atau *leaflet*, *wachart*, foto/gambar, ataupun model/maket (Kementerian Pendidikan dan Kebudayaan, 2010, p. 27). Bahan ajar berfungsi sebagai pedoman bagi guru untuk mengarahkan aktivitas selama proses kegiatan belajar mengajar berlangsung. Buku pelajaran adalah salah satu sumber belajar yang memiliki peran cukup besar dalam proses belajar siswa. Sitepu (2005, p. 116) menyatakan bahwa buku teks pelajaran juga memiliki kemampuan sebagai pendorong peningkatan mutu proses maupun hasil pembelajaran. Berdasarkan hasil prasarvei diperoleh fakta bahwa di Madrasah Negeri Kabupaten Magetan belum menggunakan modul sebagai bahan ajar, padahal hasil wawancara dengan siswa, diperoleh fakta bahwa mereka senang dan merasa termotivasi ketika dalam kegiatan belajar mengajar menggunakan modul.

Seorang guru dituntut untuk merencanakan apa yang akan dilakukan dalam mengelola kegiatan pembelajaran. Penyusunan Rencana Pelaksanaan Pembelajaran (RPP) sangatlah penting karena RPP mencakup desain pembelajaran yang dibuat dan dirancang, untuk selanjutnya dilaksanakan dalam kegiatan belajar mengajar untuk mencapai tujuan yang diinginkan. Keberadaan RPP diharapkan akan memperlancar, mengefektifkan, dan mengoptimalkan mutu proses pembelajaran. Rancangan, perencanaan pembelajaran didesain dalam bentuk Silabus. Perencanaan pembelajaran setidaknya meliputi penyusunan rencana pelaksanaan pembelajaran, penyiapan media dan sumber belajar,

perangkat penilaian pembelajaran, dan skenario pembelajaran. (Kementerian Pendidikan dan Kebudayaan, 2016, p. 5). RPP dikembangkan dari silabus untuk mengarahkan kegiatan pembelajaran peserta didik dalam upaya mencapai Kompetensi Dasar (KD). Setiap pendidik diharapkan mampu menyusun RPP secara lengkap dan sistematis agar terlaksana pembelajaran yang interaktif, inspiratif, menantang, menyenangkan, efisien dan memotivasi siswa berpartisipasi aktif, dan memberikan ruang yang cukup bagi prakarsa, kreativitas dan kemandirian sesuai bakat, minat, perkembangan fisik dan psikologis peserta didik (Kementerian Pendidikan dan Kebudayaan, 2016, p. 6).

Keterampilan berpikir kritis peserta didik dapat dikembangkan melalui model pembelajaran yang tepat. Di antara berbagai macam model pembelajaran yang dipercaya dapat melatih keterampilan berpikir kritis untuk siswa, adalah pembelajaran berbasis konstruktivistik (Thompson, 2011, p. 2). Vygotsky (Vijayaratnam 2009, p. 56) menyatakan dalam pembelajaran konstruktivistik siswa diberikan kesempatan untuk bereksperimen atau berinkuiri dan berinteraksi dengan teman sejawatnya untuk meningkatkan perkembangannya. Keterampilan berpikir kritis dapat dilatihkan melalui model pembelajaran berbasis inkuiri sesuai tuntutan Kurikulum 2013 melalui pendekatan ilmiah. Lawson (1995, p. 200) menyatakan bahwa pembelajaran biologi di sekolah menengah dalam kurikulum yang berbasis inkuiri mampu memberdayakan keterampilan berpikir kritis dan penguasaan konsep siswa.

Karakteristik RPP, modul dan instrumen berbasis *inquiry lesson* adalah memuat sintaks pembelajaran *inquiry lesson* dan diintegrasikan dengan aspek berpikir kritis. Perangkat pembelajaran *inquiry lesson* yang dikembangkan oleh peneliti mengacu pada Facione (2015, p. 8) yang meliputi enam tahap yaitu; interpretasi, analisis, inferensi, evaluasi, penjelasan serta pengaturan diri yang didesain untuk memotivasi siswa belajar secara mandiri dalam menemukan prinsip, dan konsep materi pembelajaran. Sintaks *inquiry lesson* yang dimuat menurut

Wenning (2011, p. 10) yaitu; observasi (*observation*) diintegrasikan dengan aspek interpretasi, manipulasi (*manipulation*) diintegrasikan dengan aspek analisis, generalisasi (*generalization*) diintegrasikan dengan aspek inferensi, verifikasi (*verification*) diintegrasikan dengan aspek evaluasi, dan aplikasi (*aplication*) diintegrasikan dengan aspek penjelasan dan pengaturan diri. Berdasarkan apa yang telah diuraikan, maka dirasa perlu dilakukan pengembangan perangkat pembelajaran yang dapat membantu guru dan siswa dalam melaksanakan pembelajaran berbasis inkuiri untuk meningkatkan keterampilan berpikir kritis siswa.

Berdasarkan uraian yang telah dipaparkan tersebut, penelitian ini bertujuan untuk mengetahui efektivitas perangkat pembelajaran Biologi untuk meningkatkan keterampilan berpikir kritis siswa.

Metode Penelitian

Penelitian ini merupakan penelitian deskriptif kualitatif dan kuantitatif yang dilakukan untuk mengetahui efektivitas perangkat pembelajaran pada materi sistem ekskresi manusia berbasis *inquiry lesson* dalam meningkatkan keterampilan berpikir kritis siswa. Tempat penelitian dipilih dengan *purposive sampling* dengan kriteria: (a) telah terakreditasi A menurut Badan Standar Nasional Pendidikan di Indonesia; (b) madrasah negeri yang memiliki jumlah siswa terbanyak pada Tahun Pelajaran 2016/2017; (c) guru telah memiliki pengetahuan dan melaksanakan pembelajaran berbasis inkuiri. Berdasarkan kriteria tersebut, penelitian ini dilaksanakan pada Madrasah Aliyah Negeri di Kabupaten Magetan.

Selanjutnya dipilih dua kelas secara *random sampling*, satu kelas ditetapkan sebagai *existing class* dan satu kelas sebagai *experiment class*. Hasilnya diperoleh subjek dalam penelitian adalah 22 siswa pada kelas MIA 2 sebagai kelas eksperimen dan 22 siswa kelas MIA 1 sebagai *existing class*. Kelas eksperimen diberikan pembelajaran menggunakan RPP, modul dan instrumen penilaian berdasarkan indikator berpikir kritis sedangkan pada kelas kontrol kegiatan belajar mengajar

berlangsung seperti biasanya tanpa RPP, modul dan instrumen yang dikembangkan berdasarkan indikator berpikir kritis.

Data penelitian diperoleh melalui observasi, wawancara, angket, telaah dan tes. Observasi dilakukan untuk mencermati keterlaksanaan pembelajaran berbasis inkuiri pada proses belajar mengajar oleh guru di dalam kelas. Wawancara kepada guru dilakukan untuk mengetahui pengetahuan dan pemahaman guru terhadap pembelajaran inkuiri dan kebutuhan bahan ajar. Selain itu, Wawancara juga dilakukan untuk memperoleh informasi tentang kebutuhan bahan ajar siswa dan kesiapan siswa terhadap pembelajaran inkuiri. Kuisisioner diberikan kepada siswa untuk mengetahui level inkuiri siswa. Tes kepada siswa dilakukan untuk mengetahui profil awal keterampilan berpikir kritis siswa, dan hasil keterampilan berpikir kritis siswa setelah menggunakan perangkat pembelajaran berbasis *inquiry lesson*, sedangkan tes kepada guru dilakukan untuk kemampuan guru dalam mengurutkan sintaks *inquiry lesson*.

Teknik analisis data yang digunakan dalam penelitian ini adalah teknik analisis deskriptif kualitatif dan kuantitatif. Teknik analisis kuantitatif digunakan untuk menghitung efektivitas perangkat pembelajaran, sedangkan analisis deskriptif kualitatif digunakan untuk menganalisis data tentang pengetahuan dan pemahaman guru terhadap pembelajaran inkuiri. Sebelum diterapkan di lapangan, perangkat pembelajaran telah melalui proses validasi yang dilakukan menggunakan formula Aiken's (*Experts Judgement*) dan diestimasi reliabilitasnya dengan koefisien Kappa. Selanjutnya, untuk mengetahui efektivitas perangkat pembelajaran dilakukan teknik analisis *quasi experimental group* berupa *existing class* dan *experiment class* yang dikembangkan dengan menggunakan desain *Pre-test-post-test Control Group Design* sebagaimana disajikan pada Tabel 1.

Tabel 1. Desain Penelitian

<i>Eksperiment Class</i>	O	X ₁	O ₁
<i>Existing Class</i>	O	X ₂	O ₁

- X1 = Perlakuan, berupa pembelajaran menggunakan RPP, modul dan instrumen berbasis *inquiry lesson*
 X2 = Pembelajaran seperti biasa yang dilakukan tanpa menggunakan RPP, modul dan instrumen berbasis *inquiry lesson*
 O = *Pre-test* (tes awal)
 O₁ = *Post-test* (tes akhir)

Uji efektivitas tersebut dilakukan dengan menggunakan uji beda *Independent T-Test* terhadap hasil pretes dan postes siswa untuk mengetahui perbedaan hasil pretes dan postes siswa. Tingkat efektivitas perangkat pembelajaran diuji dengan membandingkan *gain score* ternormalisasi ($\langle g \rangle$) pada hasil *pretest* dan *posttest* baik pada *existing class* maupun *experiment class*. Hasil perhitungan *gain score* disajikan dalam Tabel 2.

Tabel 2. Kriteria Indeks *N-gain* dan Interpretasinya

<i>N-Gain</i>	Interpretasi
$0,7 < g < 1$	Tinggi
$0,3 \leq g \leq 0,7$	Sedang
$0 < g < 0,3$	Rendah

(Sumber: Hake, 1998, p. 65).

Hasil Penelitian

Pengembangan Perangkat Pembelajaran

Perangkat pembelajaran berbasis *inquiry lesson* dikembangkan berdasarkan sintaks *inquiry lesson* yang terdiri dari (1) *observation*; pada tahap ini siswa diajak untuk mengamati fenomena yang disajikan, melibatkan siswa untuk memunculkan respon identifikasi permasalahan yang terdapat dalam fenomena tersaji; (2) *manipulation*; pada tahap ini siswa mengidentifikasi faktor-faktor yang mempengaruhi fenomena dan mengusulkan ide-ide yang memungkinkan untuk dilakukan penyelidikan; (3) *generalization*; pada tahap ini, siswa menggeneralisasi atau membuat kesimpulan atas dasar hasil temuannya; (4) *verification*; pada tahap ini siswa membuat prediksi dilanjutkan melakukan pengujian pada permasalahan lain yang serupa berda-

sarkan konsep yang diperoleh dari tahap sebelumnya selanjutnya mendiskusikannya kembali; (5) dan *application*; pada tahap ini guru memberi penguatan dan motivasi kepada siswa agar mampu menghasilkan penelitian kualitatif lainnya menggunakan media yang berbeda (Wenning, 2010, p. 9).

Karakteristik perangkat pembelajaran berbasis *inquiry lesson* dibandingkan perangkat lain adalah perangkat pembelajaran ini mengintegrasikan sintaks *inquiry lesson* dengan aspek berpikir kritis sehingga diharapkan siswa lebih terlibat aktif dalam menemukan konsep secara mandiri. Dalam model pembelajaran *inquiry* siswa didorong belajar tentang materi sains dan lebih memahami sifat-sifat sains (Zion, 2007, p. 424). Hal ini relevan dengan hakekat teori konstruktivisme, yang mensyaratkan siswa secara aktif terlibat untuk memperoleh pengertian yang ingin diketahuinya (Suparno, 1997, p. 49).

Uji validitas kelayakan perangkat pembelajaran berbasis *inquiry lesson* dilakukan melalui (1) uji coba lapangan awal yang melibatkan validator ahli materi, ahli pengembangan modul, ahli perangkat pembelajaran, praktisi (guru), dan 12 siswa (2) uji coba lapangan utama; dan (3) uji coba lapangan operasional.

Validasi telaah RPP, modul dan instrumen penilaian Hasil perhitungan formula Aiken's menunjukkan bahwa lembar telaah RPP memiliki koefisien Aiken's V sebesar 0,93, modul sebesar 0,92 dan butir instrumen penilaian sebesar 0,90. Hal ini berarti bahwa RPP, modul, dan butir instrumen penilaian memiliki validitas isi yang baik. (Aiken, 1985, p. 133).

Uji reliabilitas perangkat RPP menggunakan koefisien Kappa diperoleh hasil sebesar 0,659. Uji reliabilitas *interrater* pada modul diperoleh koefisien Kappa sebesar 0,686. Adapun Hasil analisis reliabilitas butir soal diperoleh Cronbach's alpha sebesar $0,757 > 0,423$ yang berarti semua butir soal reliabel.

Berdasarkan hasil observasi proses pembelajaran di kelas pada Madrasah tempat penelitian diperoleh fakta bahwa guru sudah melaksanakan pembelajaran berbasis

inkuiri, namun dalam pelaksanaan, sintaks yang diterapkan belum lengkap dan masih didominasi *discovery learning*. Hal ini diperkuat dengan hasil angket level inkuiri siswa yang memperoleh data bahwa rata-rata keterampilan *Discovery Learning* sebesar 65,46%, *Interactive Demonsration* sebesar 62,35%, *Inquiry lesson* sebesar 51,82%, *Guided Inquiry Laboratory* sebesar 42,50% dan *Bounded Inquiry Laboratory* sebesar 30,42%.

Berdasarkan hasil wawancara dan angket diketahui bahwa persepsi guru terhadap pembelajaran berbasis inkuiri dalam kategori cukup, dengan perolehan rata-rata 66,25%. Sedangkan kemampuan rata-rata mengurutkan sintaks inkuiri menunjukkan hasil sebesar 50,35% dan dalam kategori kurang, dengan rincian sebagai berikut; persentase kemampuan mengurutkan sintaks pembelajaran inkuiri pada tahap pendahuluan sebesar 50,83% tergolong kategori kurang; tahap inti sebesar 68,80% tergolong kategori cukup; dan pada tahap penutup sebesar 33,33% dan tergolong kategori sangat kurang.

Hasil telaah bahan ajar berupa buku Biologi yang digunakan sebagai rujukan utama bagi guru dalam kegiatan pembelajaran terhadap pemberdayaan berpikir kritis diperoleh rata-rata sebesar 39,90% dan tergolong dalam kategori kurang. Berdasarkan aspek kemampuan berpikir kritis siswa menunjukkan bahwa: (1) Aspek interpretasi rata-rata sebesar 47,60%; (2) aspek analisis rata-rata sebesar 44,79%; (3) aspek evaluasi rata-rata 33,89; 4) aspek kesimpulan diperoleh rata-rata 40,82%; (5) aspek penjelasan diperoleh rata-rata 45,44%; (6) aspek pengaturan diri diperoleh rata-rata 26,90%.

Setiap kegiatan dalam RPP, Modul dan instrumen penilaian berisi aktivitas terstruktur yang tersusun atas sintaks pembelajaran berbasis *inquiry lesson* yang diintegrasikan dengan aspek berpikir kritis. Tahap observasi diintegrasikan dengan aspek interpretasi, sedangkan tahap manipulasi diintegrasikan dengan aspek analisis. Selama ini guru belum pernah menyusun modul pembelajaran secara mandiri. Bahan ajar yang digunakan bersumber dari penerbit. Bahkan bahan ajar terbitan beberapa tahun lalu yang

berbasis KTSP pun masih digunakan dengan alasan materi yang disajikan masih relevan dengan Kurikulum 2013.

Hasil wawancara dengan siswa diketahui bahwa siswa merasa lebih bersemangat belajar menggunakan modul daripada LKS yang selama ini mereka gunakan. Hal ini dikarenakan modul memiliki gambar yang lebih jelas dan besar, serta warna yang lebih menarik. Tahapan dalam modul juga membuat siswa lebih mudah untuk memahami konsep materi yang sedang dipelajari.

Perangkat pembelajaran dikembangkan dengan menggunakan lima langkah sintaks pembelajaran *Inquiry lesson* yang diintegrasikan dengan enam aspek berpikir kritis. Materi sistem ekskresi manusia dibagi menjadi tiga kegiatan, yaitu sistem ekskresi pada ginjal, sistem ekskresi pada kulit dan sistem ekskresi pada hati dan paru-paru.

Tahap generalisasi diintegrasikan dengan aspek inferensi, tahap verifikasi diintegrasikan dengan aspek evaluasi dan tahap aplikasi diintegrasikan dengan aspek penjelasan dan pengaturan diri.

Tabel 3. Hasil Analisis Deskriptif Skor *Pretest* dan *Posstest*

Jenis Tes	Kelas	Σ	Mean	Std	Min	Maks
Pretes	<i>existing</i>	22	46,6	6,97	32,5	67,50
	<i>Exsperiment</i>	22	48,2	6,34	42,5	67,50
Rata-Rata			47,5	6,65	32,5	67,50
Postes	<i>existing</i>	22	60,0	5,61	32,8	77,50
	<i>Exsperiment</i>	22	80,1	6,38	60,0	90,00
Rata-Rata			70,1	5,99	46,4	83,75

*Skor ideal maksimal adalah 100

Tabel 3 menunjukkan hasil analisis deskripsi kemampuan berpikir kritis siswa sebelum diberikan perlakuan dan setelah dilakukan perlakuan pada *existing class* (pembelajaran biasa) dan *experiment class* (menggunakan perangkat pembelajaran berbasis *inquiry lesson*). Rata-rata skor *pre test* kelas kontrol adalah sebesar 46,6; dengan standar deviasi 6,97; skor minimum 32,5; dan maksimum 67,5. Rata-rata skor pada *pretest* kelas eksperimen sebesar 48,2; dengan standar deviasi 6,34; skor minimum 42,5 dan maksimum 67,50. Rata-rata *post test* kelas kontrol

adalah sebesar 60,07; dengan standar deviasi 5.61; skor minimum 32,78 dan skor maksimum 77,50. Rata-rata *postest* kelas eksperimen adalah sebesar 80,12; dengan standar deviasi sebesar 6,38; skor minimum 60,09 dan skor maksimum 90,00.

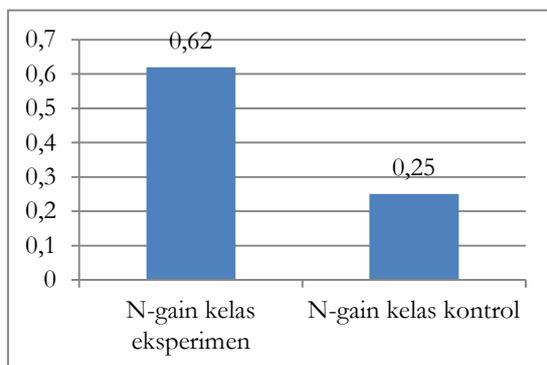
Efektivitas Perangkat Pembelajaran

Untuk mengetahui keefektifan perangkat pembelajaran berbasis *inquiry lesson* dilakukan *independent t test*. Hasil uji *independent t test* menunjukkan indeks korelasi sebesar $0,475 > 0,05$ dan *sig 2-tailed* 0,00 sebagaimana disajikan dalam Tabel 4.

Tabel 4. Analisis *T-test* pada Kelas Eksperimen (MIA 2)

Uji	Hasil	Keputusan	Kesimpulan
T test	Sig 2-tailed = 000	Sig 2 tailed < 0.05, sehingga Ho ditolak	Terdapat perbedaan antara hasil pretes dan postes

Dalam rangka mengetahui perbedaan hasil *pretest* dan *posttest* pada kelas kontrol dan kelas eksperimen dilakukan penghitungan *N-gain* skor pada *pretest* dan *posttest* sebagaimana Gambar 1.



Gambar 1. Histogram *N-gain* Kelas Eksperimen dan *Existing Class*

Dari Gambar 1 diketahui bahwa *N-gain* hasil pretes dan postes pada kelas eksperimen adalah sebesar 0,62 dalam kategori sedang, sedangkan *N-gain* pada kelas kontrol antara hasil pretes dan postes diperoleh 0,25 dalam kategori rendah. Berdasarkan hal tersebut diketahui bahwa terdapat perbedaan keterampilan berpikir kritis siswa baik

pada kelas kontrol maupun kelas eksperimen, namun perbedaan pada kelas kontrol lebih kecil dibandingkan dengan kelas eksperimen.

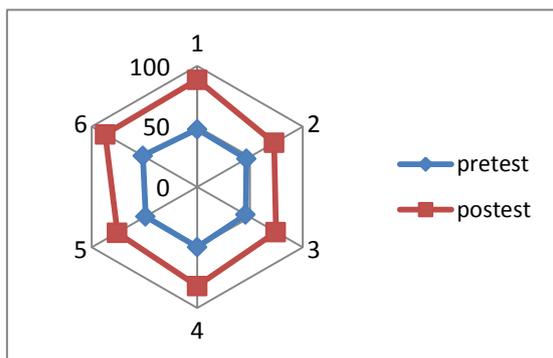
Pembahasan

Perangkat pembelajaran berbasis *inquiry lesson* merupakan perangkat yang terdiri dari Rencana Pelaksanaan Pembelajaran (RPP), Modul dan Instrumen penilaian yang didesain mengacu pada sintaks pembelajaran *inquiry lesson* dan diintegrasikan dengan aspek berpikir kritis pada materi sistem ekskresi manusia. *Inquiry lesson* merupakan salah satu tingkatan dari model pembelajaran *level of inquiry*, menurut Wenning (2010, p. 4) pembimbingan secara perlahan dilepaskan sehingga siswa dapat melakukan *inquiry* secara mandiri, namun guru tetap memberikan bimbingan dan mendukung dengan berbagai pertanyaan untuk membantu siswa melaksanakan proses penyelidikan.

Pada tahap observasi pelaksanaan pembelajaran menggunakan perangkat pembelajaran pada kelas eksperimen pada pertemuan ke-1 menunjukkan bahwa siswa sangat antusias dalam mengikuti pembelajaran, meskipun siswa belum memahami perbedaan membuat identifikasi masalah dan merumuskan masalah. Pada tahap guru belum maksimal dalam memberikan motivasi. Pada pertemuan ke-2 siswa telah lebih memahami perbedaan antara identifikasi masalah dan rumusan masalah dan guru lebih tepat dalam memberikan motivasi kepada siswa.

Efektivitas perangkat pembelajaran ditunjukkan dengan adanya peningkatan keterampilan berpikir kritis siswa setelah dianalisis berdasarkan *score N-Gain*. Berdasarkan Gambar 1 diketahui rata-rata *N-gain* keterampilan berpikir kritis siswa *existing class* sebesar 0,25 atau termasuk dalam kategori rendah. Pada *experiment class* diperoleh *N-gain* sebesar 0,62 atau termasuk dalam kategori sedang. Peningkatan *N-gain* yang signifikan pada kelas eksperimen tersebut merupakan akibat dari penggunaan perangkat pembelajaran berbasis *inquiry lesson* yang mampu melatih dan meningkatkan keterampilan

pilan berpikir kritis siswa. Dalam desain RPP, modul dan instrumen penilaian disajikan fenomena ataupun permasalahan nyata yang ada di lingkungan siswa sehingga membantu siswa dalam menemukan konsep pengetahuan melalui proses penemuan. Pencapaian persentase keterampilan berpikir kritis masing-masing aspek indikator *pretest* dan *posttest* disajikan dalam Gambar 2.



Gambar 2. Persentase Pencapaian Indikator Berpikir Kritis Siswa

- | | |
|-------------------|---------------------|
| 1 : Interpretasi; | 2 : Analisis; |
| 3 : Inferensi; | 4 : Evaluasi; |
| 5 : Penjelasan; | 6 : Pengaturan Diri |

Hasil *pretest* rata-rata persentase keterampilan berpikir kritis siswa sebesar 52,28% tergolong kategori kurang. Persentase pada aspek interpretasi sebesar 48,80% dalam kategori sangat kurang, aspek analisis sebesar 45,98% dalam kategori sangat kurang, aspek evaluasi sebesar 53,39% kategori sangat kurang, aspek kesimpulan sebesar 55,09% dalam kategori kurang; aspek penjelasan sebesar 46,48% kategori sangat kurang, dan aspek pengetahuan diri sebesar 63,94% kategori cukup. Sedangkan pada hasil *posttest* diperoleh kenaikan rata-rata persentase 31,93% sehingga menjadi 80,12%. Aspek interpretasi diperoleh rata-rata 88,67% kategori sangat baik, aspek analisis sebesar 72,92% kategori cukup, aspek inferensi sebesar 74,50% kategori cukup, aspek evaluasi sebesar 81,83% kategori baik, aspek penjelasan sebesar 75,75% kategori cukup dan aspek pengaturan diri sebesar 87,08% tergolong kategori baik.

Hal ini relevan dengan yang dikemukakan dalam National Research Council

(1996, p. 13) yang menyatakan bahwa dalam proses pembelajaran guru sebagai bagian dari komunitas pembelajaran didalam kelas harus mampu menyusun perencanaan pembelajaran, membuat keputusan tentang apa yang akan dipelajari siswa, bagaimana cara mereka belajar, dan bagaimana sumberdaya dialokasikan secara berkelanjutan sehingga siswa fokus pada pembelajaran secara aktif bukan hanya sekedar hafalan, yang pada akhirnya terbangun basis pengetahuan dan pemahaman konsep. Dalam pembelajaran berbasis inkuiri siswa dilibatkan dalam penelitian, investigasi, mereka berinteraksi dengan guru dan teman sejawat. Siswa menerapkan pengetahuan mereka dan terlibat dalam pemecahan masalah, perencanaan, membuat keputusan, diskusi kelompok dan mereka mengalami penilaian yang konsisten dengan pendekatan aktif selama proses pembelajaran. Minner, Levy, & Century (2010, p. 475) juga menyatakan pembelajaran *inquiry* dipercaya sebagai salah satu model konstruktivisme yang didalamnya membangun proses pengetahuan siswa melalui penyelidikan fenomena ilmiah. Dalam pendekatan *inquiry lesson* siswa dilibatkan secara kelompok maupun individu dalam kegiatan inkuiri. Hal ini diharapkan terjadinya kerja sama dan saling membantu antarsiswa dalam belajar. Josef, Trnova, & Sibor (2012, p. 208) memperoleh fakta berdasarkan penelitiannya bahwa pembelajaran sains berbasis *inquiry* berhasil diterapkan sebagai salah satu metode pembelajaran karena mendorong motivasi siswa dalam belajar. Hal ini relevan dengan hasil pengamatan dan wawancara selama kegiatan belajar mengajar berlangsung, bahwa siswa lebih termotivasi dan lebih menyukai kegiatan belajar mengajar yang dalam prosesnya melibatkan siswa secara aktif baik dalam eksperimen, maupun observasi. Demikian juga menurut teori Vygotsky bahwa fungsi mental yang lebih tinggi pada umumnya muncul dalam percakapan dan kerja sama antar-individu sebelum fungsi mental yang lebih tinggi itu terserap kedalam individu tersebut (Trianto, 2010, p. 76).

White, *et al.* (2009, p. 30) mengemukakan bahwa keterampilan berpikir kritis

siswa dapat ditingkatkan dengan mengimplementasikan pembelajaran yang berpusat pada siswa (*student centered learning*). Dalam kegiatan pembelajaran berbasis *inquiry lesson* siswa diarahkan untuk melakukan percobaan ilmiah dan bertindak layaknya seorang peneliti dalam *setting* penelitian formal. Siswa dalam hal ini harus dapat mendefinisikan sistem, menentukan variabel dependen dan independen yang melibatkan variabel kontrol dan manipulasi. Jadi proses pembelajaran lebih berpusat pada siswa, sedangkan guru bertindak sebagai fasilitator dan mediator (Wenning, 2011, p. 5). Siswa dituntut lebih aktif dalam menemukan konsep secara mandiri, sedangkan guru memberikan bimbingan, panduan dan membantu siswa untuk melakukan eksperimen, identifikasi, pengendalian variabel dan pendefinisian. Kasmurie, Abdul, & Ahmad, (2010, p. 272) membuktikan bahwa pembelajaran *inquiry* mampu meningkatkan keterampilan berpikir kritis siswa, sehingga perlu ditekankan pada setiap sekolah sebagai salah satu alternatif metode pembelajaran di sekolah. Hal ini selaras dengan hasil penelitian Purwanto (2013, p. 110) bahwa dengan model pembelajaran *level of inquiry* dapat melatih kemampuan berinkuiri siswa yang didalamnya sejalan dengan indikator keterampilan berpikir kritis siswa.

Pembelajaran berbasis inkuiri yang diintegrasikan dengan modul memberikan hasil yang signifikan dalam meningkatkan keterampilan berpikir kritis siswa. Kumari & Kulshrestha (2013, p. 5) memperoleh hasil bahwa pembelajaran berbasis *inquiry* yang diintegrasikan dengan modul menjadikan pengalaman belajar siswa menjadi jauh lebih bermakna dan lebih terlibat aktif dalam kegiatan belajar mengajar. Perangkat pembelajaran berbasis *inquiry lesson* merupakan perangkat yang didesain dengan memuat permasalahan yang terjadi di lingkungan yang memerlukan pembuktian proses eksperimen melalui metodologi ilmiah yang menstimulasi siswa untuk menemukan konsep secara mandiri sehingga menghasilkan penemuan yang bermakna. Hal tersebut relevan dengan teori belajar Ausubel bahwa berusaha

sendiri untuk mencari pemecahan masalah serta pengetahuan yang menyertainya sehingga menghasilkan pengetahuan yang benar-benar bermakna (Dahar, & Wilis, 2011, p. 134). Selaras dengan pernyataan Buffington (2007, p. 22) bahwa keterampilan berpikir kritis siswa dapat berkembang apabila dalam kegiatan belajar mengajar secara periodik guru melibatkan keterampilan berpikir kritis dalam tahapannya. Hal senada diupayakan dalam perangkat pembelajaran berbasis *inquiry lesson*, dalam RPP, dan modul berisikan kegiatan pembelajaran secara periodik memuat keterampilan siswa untuk mengobservasi, menginterpretasi data, menganalisis, menjelaskan. L. M. Sartorelli dan R. Swartz (Hossoubah, 2007, pp. 96-110) menyatakan bahwa ada beberapa cara dalam usaha peningkatan keterampilan berpikir kritis siswa diantaranya: (1) membaca secara kritis; (2) mengembangkan daya analisis pada suatu permasalahan melalui kegiatan diskusi untuk mengatasi permasalahan tersebut, mencari solusi terbaik dan mempertimbangkan dampak positif dan negatifnya; (3) mengembangkan kemampuan mengamati atau observasi, mengidentifikasi kelebihan dan kekurangannya, pro-kontra dari permasalahan tersebut; (4) meningkatkan rasa ingin tahu siswa, menggali kemampuan bertanya dan refleksi, mengajukan pertanyaan berkualitas yang tidak hanya membutuhkan jawaban benar dan salah saja. Dalam perangkat pembelajaran berbasis *inquiry lesson* secara berurutan disajikan fenomena untuk kegiatan observasi melatih keterampilan interpretasi siswa, tahap manipulasi untuk melatih analisis siswa, tahap generalisasi melatih keterampilan siswa dalam menyusun kesimpulan, tahap verifikasi melatih evaluasi siswa, dan tahap aplikasi untuk melatih keterampilan penjelasan dan pengaturan diri siswa sehingga pada akhirnya keterampilan berpikir kritis siswa secara umum dapat meningkat.

Simpulan

Kesimpulan dari penelitian ini adalah perangkat pembelajaran berbasis *inquiry lesson* efektif dalam meningkatkan keterampilan

pilan berpikir kritis siswa pada materi sistem ekskresi manusia. Meskipun pada pertemuan pertama siswa merasa ada beberapa hambatan yang dialami, diantaranya siswa masih belum terbiasa dalam menyusun rumusan masalah dan hipotesis. Namun hal tersebut sudah berjalan baik pada pertemuan selanjutnya. Hasil uji statistik diperoleh perbedaan signifikan antara hasil *pretest* dengan hasil *posttest* pada kelas eksperimen dengan kategori sedang. Sedangkan pada *existing class* menunjukkan perbedaan hasil *pretest* dengan *posttest* yang lebih rendah dibandingkan dengan *gap* pada kelas eksperimen dengan kategori rendah.

Adapun saran bagi peneliti lanjut, hasil penelitian ini dapat digunakan sebagai dasar untuk melakukan penelitian terkait dengan proses berpikir kritis siswa khususnya siswa yang memiliki kesamaan kondisi dengan siswa di Madrasah tempat penelitian ini dilakukan. Selain itu, perlu dilakukan uji coba lebih lanjut untuk memperoleh data yang akurat tentang perkembangan berpikir kritis siswa dengan mengacu pada indikator berpikir kritis menurut ahli yang lain seperti Ennis, Paul dan lainnya dan dilaksanakan pada bidang pelajaran yang berbeda.

DAFTAR PUSTAKA

- Aiken, L. R. (1985). Three coefficients for analyzing the reliability and validity of rating. *Educational and Psychological Measurement*, 45, 131-142.
- Balitbang. (2013). *Survey internasional PISA (Programme for International Student Assessment student)*. Jakarta: Kementerian Pendidikan dan Kebudayaan. Retrieved November 17, 2017 from <http://litbang.kemdikbud.go.id/index.php/survey-internasional-pisa>
- Buffington, M. L. (2007) Contemporary approaches to critical thinking and the world wide web. *Art Education*, 60(1), 18-23.
- Dahar, D. & Wilis, R. (2011). *Teori-teori belajar dan pembelajaran*. Jakarta: Erlangga.
- Demirhan, E., & Koklukaya, A. N. (2013). The critical thinking dispositions of prospective science teachers. In *5th world Conference on Educational Sciences – WCES 2013. Procedia Social and Behavioral Sciences*, 116, 1551-1555.
- Facione, P.A. (2015). *Critical thinking : what it is and Why it Count*. Retrieved December 17, 2017 from <http://www.insightassessment.com/content/download/1176/7580/file/what/26why2010.pdf>
- Friedel, C., Irani, T., Rudd, R., Gallo, M., Eckhardt, E., & Ricketts, J. (2008). Overtly teaching critical thinking and inquiry-based learning: A comparison of two undergraduate biotechnology classes. *Journal of Agricultural Education*, 49(1), 72-84. Retrieved from <http://step.ufl.edu/resources/criticalthinking/ctmanual.pdf>
- Hake, R. (1998). Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *American Journal of Physics*, 66(1), 64-74.
- Hossoubah, Z. (2007). *Developing creative and critical thinking skills* (Terjemahan). Bandung: Yayasan Nuansa Cendekia.
- Josef, T., Trnova, E., & Sibor, J. (2012). Implementation of inquiry based science education in science teacher training. *International Journal On New Trend In Education And Their Implication*, 2, 199-20.
- Kasmuric, A. K, Abdul, R. A., & Ahmad, A. S. (2010). The effectiveness of inquiry teaching in enhancing students critical thinking. *Procedia Social and Behavioral Sciences*, 7(C), 264-27.
- Kementerian Pendidikan dan Kebudayaan. (2010). *Juknis pengembangan bahan ajar SMA*. Jakarta: Kementerian Pendidikan dan Kebudayaan Republik Indonesia.

- Kumari, K., & Kulshrestha, A. K. (2013). Impact of constructivist inquiry-based learning approach on science achievement at grade VIII. *International Journal of Applied and Studies*, 2, 1-5.
- Lawson, A. E. (1995). *Science teaching and the development of thinking*. California: Wadsworth Publishing Company.
- Minner, D. D., Levy, A. J., & Century, J. (2010). Inquiry-based science instruction—What is it and does it matter. results from a research synthesis Years 1984 to 2002. *Journal of Research in Science Teaching*, 47, 474-494.
- National Center for Educational Statistics. (2015). *Trends in international mathematics and science study (TIMSS): Science for grades 4 and 8: averages*. Retrieved December 15, 2017, from <https://nces.ed.gov/timss/>,
- National Research Council. (1996). *National science education standards*. Washington, D.C.: National Academies Press.
- Norris, S. P., & Ennis, R. H. (1989). *Evaluating critical thinking*. Critical Thinking Press & Software.
- Kementerian Pendidikan dan Kebudayaan. (2016) *Permendikbud Nomor 22 Tahun 2016 Tentang Standar Proses Pendidikan Dasar dan Menengah*. (Jakarta. Kemdikbud Republik Indonesia
- Purwanto. (2013). Analisis kemampuan inkuiri dan hasil belajar siswa sekolah menengah pertama melalui model pembelajaran berbasis model hierarki of inquiry. In *Prosiding Pertemuan Ilmiah XXVII HFI Jateng & DIY*, 107-110.
- Sitepu, M. A. (2005). Memilih buku pelajaran. *Jurnal Pendidikan Penabur*, 4, 113-126.
- Suparno, P. (1997). *Filsafat konstruktivisme dalam pendidikan*. Yogyakarta: Kanisius.
- Sternberg, R. J. (1986). Critical thinking it's measurement, and improvement. *Jurnal National Inst. Education*, (ED), 1-37.
- Suratno, S., & Kurniati, D. (2017). Implementasi model pembelajaran math-science berbasis performance assessment untuk meningkatkan kemampuan berpikir kritis siswa di daerah perkebunan kopi Jember. *Jurnal Penelitian Dan Evaluasi Pendidikan*, 21(1), 1-10. <https://doi.org/10.21831/pep.v21i1.11799>
- The Trends In International Maths and Science Study. (2015). *World Pupil Rankings In Science And Maths 2015*. Retrieved December 5, 2017 from www.telegraph.co.uk,
- The Partnership for 21st Century Skills (2015). *A Framework for 21st century learning*. Retrieved October 1, 2017, from www.p21.org
- Thompson, C. (2011). Critical across the curriculum: process over output. *International Journal of Humanities and Social Science*, 1(9), 1-7.
- Trianto. 2010. *Mendesain model pembelajaran inovatif progresif: konsep, landasan, dan implementasinya pada kurikulum tingkat satuan pendidikan (KTSP)*. Jakarta: Kencana.
- Vijayaratham, P. (2012). Developing higher order thinking skills and team commitment via group problem solving: A bridge to the real world. *Procedia Social and Behavioral Sciences* 66, 53-63.
- Wenning, C.J. (2010). Levels of inquiry: Using inquiry spectrum learning sequences to teach science. *Journal of Physics Teacher Education Online*, 5(4), 11-19
- Wenning, C. J. (2011). The levels of inquiry model of science teaching. *Journal of Physics Teacher Education*, 6 (2), 9-16.
- White, T. K., Whitaker, P., Gonya, T., Hein, R., Kroening, D., Lee, K., Lee, L., Lukowiak, A., & Hayes, E. (2011). The use of interrupted case studies to enhance critical thinking skills in Biology. *Journal of MicroBiology and*

Biology Education, 10, 25-31.

Prihatni, Y., Kumaidi, K., & Mundilarto, M. (2016). Pengembangan instrumen diagnostik kognitif pada mata pelajaran IPA di SMP. *Jurnal Penelitian Dan Evaluasi Pendidikan*, 20(1), 111-125. <https://doi.org/10.21831/pep.v20i1.7524>

Zion. (2007). The Spectrum of dynamic

inquiry teaching practices. *Research Science Education*, 37, 423-447.

Zikovic, Z.(2016). A model of critical thinking as an important attribute for success in the 21st century. In *International Conference on Teaching and Learning English as an Additional Language, GlobELT 2016. Procedia Social and Behavioral Science*, 232 (pp. 102-108).

EVALUASI PELAKSANAAN STANDAR PROSES PENDIDIKAN PADA SMP NEGERI DI KABUPATEN SLEMAN

Lantip Diat Prasajo^{1*}, *Fredrik Abia Kande*², *Amirul Mukminin*³

¹Universitas Negeri Yogyakarta, ²Universitas Tribuana Kalabahi, ³Universitas Jambi

¹Jl. Colombo No. 1, Depok, Sleman 55281, Yogyakarta, Indonesia

²Welai Tim., Tlk. Mutiara, Kabupaten Alor, Nusa Tenggara Timur, Indonesia

³Mendalo Darat, Jambi Luar Kota, Kota Jambi, Jambi 36122, Indonesia

* Corresponding Author. Email: lantip@uny.ac.id

Abstrak

Penelitian ini merupakan penelitian evaluasi yang bertujuan untuk mengungkapkan tingkat keefektifan pelaksanaan standar proses pendidikan. Model evaluasi yang digunakan adalah *discrepancy evaluation model*, *Provus*. Sampel penelitian ini adalah kepala sekolah dan guru di SMP Negeri Kabupaten Sleman. Pengumpulan data dilakukan dengan menggunakan angket, wawancara, dan dokumentasi. Berdasarkan analisis kuantitatif yang dilakukan, diketahui implementasi standar proses pada aspek perencanaan pembelajaran, pelaksanaan, penilaian, dan pengawasan termasuk dalam kategori sangat efektif. Lebih lanjut diketahui bahwa implementasi standar di sekolah didukung oleh lingkungan kelas yang kondusif, peserta didik kooperatif, dan peran kepala sekolah. Inovasi media pembelajaran, proses memotivasi peserta didik, perolehan informasi baru tentang materi tambahan, dan pengidentifikasian kemampuan peserta didik merupakan faktor yang menghambat pelaksanaan standar proses.

Kata kunci: *pelaksanaan standar proses, pendidikan*

AN EVALUATION OF THE IMPLEMENTATION OF EDUCATION PROCESS STANDARD ON JUNIOR HIGH SCHOOL IN SLEMAN REGENCY

Abstract

This study is an evaluation research which is aimed at revealing the effectiveness level of the implementation of educational process standard. The evaluation model used in this research was discrepancy evaluation model, *Provus*. The samples of the research were the principals and teachers in state junior high schools in Sleman Regency. The data were collected using questionnaires, interviews, and documentation. Based on the quantitative analysis which had been conducted, the implementation of process standard on the aspects of planning, implementing, assessing, and supervising the learning process is considered to be very effective. Further, it is also revealed that the implementation of standard at school is supported by the conducive class environment, cooperative students, and the principal's role. Learning media innovation, the process in motivating the learners, the acquisition of new information on additional material, and the identification process of the learners' ability are considered as the inhibiting factors of the implementation of the process standard.

Keywords: *implementation of process standard, education*

Permalink/DOI: <http://dx.doi.org/10.21831/pep.v22i1.19018>

Pendahuluan

Pemerintah Indonesia saat ini terus mengupayakan peningkatan mutu pendidikan. Data dari *The Learning Curve Pearson 2014* (Lestarini, 2014), menunjukkan bahwa mutu pendidikan Indonesia berada di posisi paling bawah atau ranking 40. Lembaga survei internasional lainnya, PISA (*Programme for International Student Assessment*) mensurvei 61 negara dan menempatkan Indonesia di posisi 60. Walaupun sering memenangkan berbagai olimpiade internasional dalam bidang akademik maupun nonakademik, namun secara keseluruhan pendidikan di Indonesia masih jauh dari kata berkualitas.

Pemerintah Indonesia sesungguhnya telah menetapkan standar yang harus dicapai oleh sekolah untuk menjamin pelaksanaan di lapangan, yang meliputi Standar kompetensi lulusan (SKL), standar isi, standar proses, standar pendidik dan tenaga kependidikan, standar pengelolaan, standar sarana dan prasarana, standar pembiayaan, dan standar penilaian. Pada tingkat pelaksanaannya sekolah harus merencanakan dan melaksanakan standar-standar tersebut agar mutu pendidikan dapat ditingkatkan.

Salah satu bukti nyata direncanakan dan dilaksanakannya standar-standar tersebut dapat dilihat dari pelaksanaan proses pembelajaran. Kegiatan ini merupakan inti dari pelaksanaan pendidikan di sekolah. Oleh karena itu, kegiatan pembelajaran harus dikelola dengan baik oleh pelaku pendidikan terutama guru. Pengelolaan dalam kegiatan pembelajaran disebut dengan manajemen pembelajaran.

Menurut Usman (2014, p. 7), tujuan dan manfaat manajemen pembelajaran adalah (1) terciptanya suasana belajar dan proses pembelajaran yang aktif, inovatif, efektif, kreatif, dan menyenangkan, (2) terciptanya peserta didik yang aktif mengembangkan potensi dirinya untuk memiliki kekuatan spiritual keagamaan, pengendalian diri, kepribadian, kecerdasan, akhlak mulia, serta ketrampilan yang diperlukan dirinya untuk masyarakat, bangsa, dan negara dan (3) terpenuhinya salah satu dari 4 kompetensi guru.

Begitu pentingnya pengelolaan pembelajaran, maka tenaga pendidik perlu memiliki pemahaman yang baik mengenai konsep, prinsip, maupun prosedur pembelajaran. Di samping itu pengenalan tentang karakteristik dan tipologi peserta didik secara baik akan membantu tenaga pendidik mewujudkan suasana belajar mengajar yang PAIKEM (pembelajaran aktif, inovatif, kreatif, efektif, dan menyenangkan) serta sesuai dengan perkembangan dan karakteristik kebutuhan peserta didik.

Kebutuhan peserta didik sangat bervariasi dan menyangkut kondisi sosial-budaya. Hal ini terjadi karena setiap peserta didik memiliki latar belakang sosial budaya, ekonomi, kondisi geografis, dan kebiasaan yang berbeda-beda. Berbagai perbedaan ini menuntut ketanggapan tenaga pendidik dalam mengembangkan kurikulum dan pembelajaran yang mampu memenuhi sebagian atau keseluruhan kebutuhan tersebut.

Provinsi yang menjadi salah satu potret dari keberagaman peserta didik adalah Provinsi Daerah Istimewa Yogyakarta (DIY). Provinsi DIY merupakan kota tujuan utama peserta didik dari berbagai daerah di Indonesia, sehingga mendapatkan predikat sebagai kota pendidikan. Predikat ini semakin bermakna jika melihat prestasi satuan pendidikan dalam beberapa tahun terakhir yang mendominasi peringkat teratas.

Hal ini akan menjadi menarik, apabila prestasi beberapa satuan pendidikan di Provinsi DIY dikaji berdasarkan standar proses. Pencapaian standar proses akan membuktikan bahwa terdapat kesesuaian antara input, proses, dan output yang dihasilkan oleh satuan pendidikan di DIY.

Berdasarkan uraian di atas, penelitian evaluasi yang bertujuan untuk mengevaluasi pelaksanaan standar proses pendidikan perlu dilaksanakan, sehingga dapat diketahui apakah prestasi sekolah-sekolah tersebut dari segi *output* didukung oleh kualitas proses yang sesuai dengan standar pendidikan. Penelitian ini dilakukan di beberapa sekolah pada jenjang SMP di Kabupaten Sleman, mengingat beberapa sekolah di kabupaten ini menduduki peringkat lima besar di Pro-

vinsi DIY. Di Kabupaten Sleman SMPN 4 Pakem dan SMPN 1 Godean memiliki prestasi dan indeks integritas UN yang tinggi meskipun sekolah ini terletak cukup jauh dari Kota Yogyakarta. Selain itu, SMPN 1 Kalasan dan SMPN 4 Depok yang memiliki integritas yang cukup tinggi dan pencapaian nilai UN termasuk lima besar terbaik dibandingkan sekolah-sekolah lainnya di Kabupaten Sleman.

Itulah sebabnya penelitian ini memilih tiga SMP Negeri di Kabupaten Sleman sebagai Sekolah Menengah Pertama Negeri berprestasi. Ketiga sekolah tersebut adalah SMPN 1 Godean, SMPN 1 Kalasan, dan SMPN 4 Depok. Pemilihan ketiga sekolah ini didasarkan pada pencapaian hasil UN dan indeks integritas yang dimiliki masing-masing sekolah.

Berdasarkan uraian di atas dan penelitian-penelitian terkait, maka tujuan penelitian ini yakni (1) untuk menilai efektivitas implementasi standar proses pendidikan dalam pembelajaran; (2) untuk menganalisis faktor pendukung dan penghambat proses pembelajaran.

Metode Penelitian

Studi ini menggunakan jenis penelitian evaluasi dengan pendekatan kuantitatif. Model evaluasi yang digunakan adalah *discrepancy evaluation model* yang dikembangkan oleh Provus (1969, p. 167) untuk mengetahui tingkat efektivitas implementasi standar proses di SMPN 1 Godean, SMPN 1 Kalasan, dan SMPN 4 Depok.

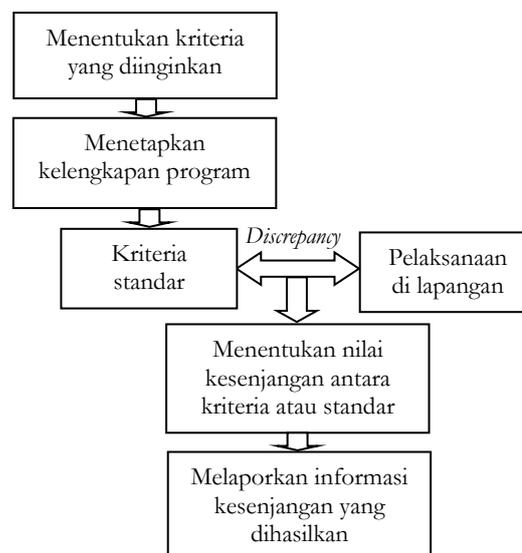
Populasi penelitian ini adalah guru-guru dari SMPN 1 Godean, SMPN 1 Kalasan, dan SMPN 4 Depok. Jumlah Populasi sampel disajikan pada Tabel 1.

Tabel 1. Populasi Penelitian

No	Nama Sekolah	Kepala Sekolah	Guru
1	SMP Negeri 1 Godean	1	28
2	SMP Negeri 1 Kalasan	1	45
3	SMP Negeri 4 Depok	1	24
	Jumlah	3	97

Berdasarkan jumlah populasi di atas, ditentukan ukuran sampel dengan menggunakan rumus dari *Issac dan Michael* (Sugiyono, 2014, p. 126). Dari rumus tersebut diperoleh sampel sebanyak 78 responden.

Langkah-langkah evaluasi *discrepancy* terhadap implementasi standar proses pendidikan dapat dilihat dalam desain berikut.



Gambar 1. Skema Tahap/Langkah Evaluasi Implementasi Standar Proses Pendidikan dengan Model Evaluasi *Discrepancy Provus*

Tahap pertama, peneliti mengidentifikasi kriteria dari standar proses dan teori pendukung evaluasi. Kedua, peneliti mengidentifikasi kelengkapan program keefektifan standar proses pendidikan SMPN 1 Godean, SMPN 1 Kalasan, dan SMPN 4 Depok. Ketiga, pelaksanaan standar proses pendidikan di SMPN 1 Godean, SMPN 1 Kalasan, dan SMPN 4 Depok dengan kriteria atau Standar Nasional Pendidikan. Keempat, mengidentifikasi kesenjangan (*discrepancy*) berupa persentase kesenjangan hasil capaian program. Kelima, peneliti memaparkan nilai kesenjangan (*discrepancy*) yang dihasilkan berdasarkan Standar Nasional Pendidikan di SMPN 1 Godean, SMPN 1 Kalasan, dan SMPN 4 Depok. Hasil evaluasi pelaksanaan Standar Nasional Pendidikan tersebut kemudian digunakan untuk mengungkap keefektifan standar proses pendidikan yang diterapkan dan kesen-

jangan antara reliabilitas di lapangan dengan kriteria keberhasilan.

Pengumpulan data pada penelitian ini dilakukan dengan menggunakan angket, wawancara dan dokumentasi. Validitas instrumen penelitian melalui validitas internal. Pengujian validitas dengan validitas konstruk (*construct validity*). Setelah aspek-aspek pada instrumen dikonstruksi, instrumen tersebut dikonsultasikan kepada ahli (*experts judgement*). Selanjutnya instrumen diujicobakan. Uji coba instrumen dilakukan di SMPN 1 Sleman dengan jumlah responden 30 guru. Responden tersebut tidak lagi terlibat dalam pengambilan data sebenarnya. Hasil pengujian validitas dan reliabilitas menunjukkan bahwa sebanyak 98 dari 123 item dinyatakan valid dan reliabel, sementara sisanya 25 item dinyatakan gugur.

Data yang terkumpul kemudian dianalisis dengan menggunakan teknik persentase. Adapun kriteria evaluasi keefektifan pelaksanaan standar proses yang digunakan dalam penelitian ini mengacu pada *Permendiknas* Nomor 41 tahun 2007 (Menteri Pendidikan Nasional Republik Indonesia, 2007a). Standar kriteria yang dimaksud meliputi: (1) Perencanaan proses pembelajaran: (a) Silabus; (b) Rencana Pelaksanaan Pembelajaran (RPP); (2) Pelaksanaan Pembelajaran; (3) Penilaian, dan (4) Pengawasan pembelajaran.

Derajat capaian standar proses diukur menggunakan pengklasifikasian kategori kriteria berdasarkan persentase capaian indikator skala empat.

Hasil Penelitian dan Pembahasan

Aspek Perencanaan Pembelajaran

Menyangkut aspek perencanaan pembelajaran, khususnya dalam penyusunan silabus dinilai dengan menggunakan 13 indikator. Total skor ketiga belas indikator dari aspek perencanaan pembelajaran dalam bentuk silabus di SMPN 1 Godean, SMPN 1 Kalasan, dan SMPN 4 Depok mencapai 94,58% dengan skor capaian 959 dari jumlah skor ideal 1.014. Selanjutnya aspek perencanaan dalam hal penyusunan rencana

pelaksanaan pembelajaran (RPP) yang dinilai dengan menggunakan enam belas indikator memperoleh skor total mencapai 97,28% dengan skor 1.214 dari skor ideal 1.248.

Pada kegiatan persiapan pelaksanaan pembelajaran diketahui bahwa total indikator dari aspek persiapan pelaksanaan pembelajaran di SMPN SMPN 1 Godean, SMPN 1 Kalasan dan SMPN 4 Depok mencapai 97,18% dengan skor 758 dari skor ideal 780. Lebih lanjut, kegiatan pendahuluan yang dinilai dari sepuluh indikator memperoleh total skor mencapai 96,15% dengan skor 750 dari skor ideal 780.

Aspek Pelaksanaan Pembelajaran

Aspek pelaksanaan pembelajaran memiliki empat belas indikator. Total skor pelaksanaan pembelajaran di SMPN SMPN 1 Godean, SMPN 1 Kalasan dan SMPN 4 Depok dari keempat belas indikator tersebut mencapai 96,98% dengan skor 1.059 dari skor ideal 1.092. Sedangkan pada kegiatan penutup dan tindak lanjut pelaksanaan pembelajaran diperoleh total skor mencapai 96,15% dengan skor 300 dari skor ideal 304 yang dinilai berdasarkan empat indikator.

Aspek Penilaian Pembelajaran

Aspek penilaian pembelajaran memiliki sembilan indikator. Total skor kesembilan indikator dari aspek pelaksanaan penilaian pembelajaran di SMPN 1 Godean, SMPN 1 Kalasan, dan SMPN 4 Depok mencapai 98,43% dengan skor 691 dari skor ideal 702. Sedangkan total skor pemanfaatan hasil penilaian yang dinilai dari empat indikator mencapai 95,83% dengan skor 299 dari skor ideal 312.

Aspek Pengawasan Pembelajaran

Aspek pengawasan di SMPN 1 Godean, SMPN 1 Kalasan, dan SMPN 4 Depok memperoleh skor capaian 12 dari skor ideal 15. Indikator pelaksanaan pengawasan memperoleh skor capaian 21 dari skor ideal 21. Indikator pengecekan pengawasan memperoleh skor capaian 18 dari skor ideal 21. Indikator tindak lanjut dan pelaporan

memperoleh skor capaian 24 dari skor ideal 24. Skor total keempat indikator dari aspek pengawasan pembelajaran di SMPN 1 Godean, SMPN 1 Kalasan, dan SMPN 4 Depok mencapai 96,15% dengan skor 75 dari jumlah skor ideal 78.

Faktor Pendukung dan Penghambat Pembelajaran

Faktor pendukung dan penghambat pembelajaran yang dapat digali melalui penelitian ini adalah: (1) guru merasa terbantu dalam pembuatan silabus dengan adanya pelatihan yang diselenggarakan pihak sekolah dengan skor capaian 72 dari skor ideal 78. (2) guru merasa terbantu dalam pembuatan RPP dengan adanya pelatihan yang diselenggarakan pihak sekolah (74 dari skor ideal 78). (3) guru merasa terbantu dalam kegiatan pembelajaran oleh fasilitas pembelajaran yang memadai dengan skor capaian 74 dari skor ideal 78. (4) guru merasa terbantu dalam kegiatan pembelajaran dengan lingkungan kelas yang kondusif (skor 77 dari skor ideal 78). (5) guru merasa terbantu dalam kegiatan pembelajaran oleh peserta didik yang kooperatif dengan skor capaian 77 dari skor ideal 78. (6) guru merasa mencapai tujuan pembelajaran karena peserta didik mampu memahami materi dengan baik (skor perolehan 73 dari skor ideal 78). (7) guru merasa terbantu dengan arahan kepala sekolah dalam menganalisis butir soal (skor capaian 66 dari skor ideal 78). (8) Indikator guru dapat mengetahui hasil belajar dengan jelas karena pelaksanaan evaluasi sesuai jadwal dengan skor 69 dari skor ideal 78.

Dengan demikian terdapat delapan aspek pendukung pembelajaran di SMPN 1 Godean, SMPN 1 Kalasan, dan SMPN 4 Depok dengan skor perolehan mencapai 93,27% dengan skor 582 dari jumlah skor ideal 624, yaitu pelatihan penyusunan silabus, pelatihan penyusunan RPP, fasilitas yang memadai, lingkungan kelas yang kondusif, peserta didik yang kooperatif, ketercapaian tujuan pembelajara, arahan kepala sekolah dalam menganalisis butir soal serta

jadwal pelaksanaan evaluasi yang tepat waktu.

Selain delapan faktor pendukung pelaksanaan standar proses pendidikan, terdapat beberapa faktor penghambat, yaitu: (1) guru merasa kesulitan dalam merancang pembelajaran aktif bagi peserta didik. (2) guru merasa kesulitan dalam menentukan metode yang tepat sesuai dengan karakteristik peserta didik. (3) guru merasa kurang mampu menginovasikan media pembelajaran. (4) guru merasa kurang mampu mengaplikasikan metode yang sesuai dengan kondisi peserta didik. (5) guru merasa kesulitan untuk memotivasi kemauan belajar peserta didik. (6) guru merasa kekurangan informasi baru (misalnya materi tambahan) yang sesuai dengan usia dan tingkatan kemampuan belajar peserta didik. (7) guru merasa kesulitan mengidentifikasi kemampuan peserta didik.

Evaluasi Pelaksanaan Pembelajaran

Hasil evaluasi pelaksanaan pembelajaran di SMPN 1 Godean, SMPN 1 Kalasan, dan SMPN 4 Depok dapat dilihat dari Tabel 2.

Tabel 2. Aspek Evaluasi Proses Pembelajaran

No Indikator	Skor	Skor Ideal
1 Silabus	94,58	100
2 RPP	97,28	100
3 Persiapan Pembelajaran	97,18	100
4 Kegiatan Pendahuluan	96,15	100
5 Kegiatan Pelaksanaan	96,98	100
6 Kegiatan Penutup	96,15	100
7 Pelaksanaan Penilaian Pembelajaran	98,43	100
8 Pemanfaatan Hasil Penilaian Pembelajaran	96,15	100
9 Tindak Lanjut Hasil Penilaian Pembelajaran	95,72	100
10 Pendukung pembelajaran	93,27	100
11 Penghambat pembelajaran	32,23	100
12 Pengawasan Pembelajaran	96,15	100
Jumlah Skor	1.090,27	1.200
Rata-rata	90,86	100
Standar Deviasi	18,51	

Selanjutnya dapat ditunjukkan capaian implementasi proses pembelajaran dari ketiga SMPN di Kabupaten Sleman berdasarkan komponen evaluasi pembelajaran standar proses pada Tabel 3.

Tabel 3. Capaian Implementasi Standar Proses Pendidikan

No	Aspek	Skor Capaian (%)	Kesenjangan
1	Perencanaan Pembelajaran		
	a. Silabus	94,58	5,42
	b. RPP	97,28	2,72
	Rata-rata	95,93	4,07
2	Pelaksanaan Pembelajaran		
	a. Persiapan Pembelajaran	97,18	2,82
	b. Kegiatan Pendahuluan	96,15	3,85
	c. Kegiatan Pelaksanaan	96,98	3,02
	d. Kegiatan Penuutup	96,15	3,85
	Rata-rata	96,62	3,38
3	Penilaian Pembelajaran		
	a. Pelaksanaan Penilaian Pembelajaran	98,43	1,57
	b. Pemanfaatan Hasil Penilaian Pembelajaran	96,15	3,85
	c. Tindak Lanjut Hasil Penilaian Pembelajaran	95,72	4,28
	Rata-rata	96,77	3,23
4	Pengawasan Pembelajaran		
	a. Kegiatan Pengawasan Pembelajaran	96,15	3,85
	Rata-rata	96,15	3,85
5	Dukungan Pembelajaran		
	Pendukung pembelajaran yang membantu guru dalam pelaksanaan pembelajaran	3,23	67,77
	Rata-rata	3,23	67,77
6	Hambatan Pembelajaran		
	a. Penghambat pembelajaran yang membantu guru dalam pelaksanaan pembelajaran	30,45	69,55
	Rata-rata	30,45	69,55
	Rata-rata total	90,86	9,14
	Standar Deviasi		18,51

Capaian implementasi standar proses pendidikan pada SMPN 1 Godean, SMPN 1 Kalasan dan SMPN 4 Depok dari kelima aspek memperlihatkan rata-rata total skor capaian 90,86 dengan kesenjangan (*discrepancy*) mencapai 9,14.

Setelah mendapatkan data hasil berupa data kuantitatif yang menunjukkan skor rata-rata adalah 90,86% dengan standar deviasi 18,51%, data tersebut dibandingkan dengan tabel kriteria yang disusun berdasarkan rerata dan standar deviasi ideal seperti yang ditunjukkan pada Tabel 4.

Tabel 4. Kriteria Evaluasi

Nilai	Interval Skor	Kriteria
A	$X < -1 SB$ $X < 52.14$	Kurang efektif
B	$X < X -1SB$ $64.11 < X \leq 52.14$	Cukup efektif
C	$X +1SB > X \geq X$ $76,08 > X \geq 64.11$	Efektif
D	$X + 1 SB$ $X \geq 76,08$	Sangat efektif

Keterangan:

X = Rerata skor ideal

$$= \frac{1}{2} (100 + 28,21) = 64.11$$

SB = Simpangan baku skor ideal

$$= \frac{1}{6} (100 - 28.21) = 11.97$$

Berdasarkan deskripsi data pada Tabel 4, maka proses pembelajaran, mulai dari perencanaan, pelaksanaan, penilaian, pengawasan, kegiatan tindak lanjut serta faktor pendukung dan penghambat pembelajaran merupakan aspek terpenting dalam penyelenggaraan pendidikan di sekolah. Oleh karena itu mutu pendidikan di sekolah dapat diukur dari, pertama, efektivitas perencanaan pembelajaran yang berkaitan dengan silabus maupun RPP merupakan salah satu tahapan penting dalam sistem pembelajaran. Sebagaimana yang dikemukakan oleh Newman (Majid, 2008, pp. 15–16) bahwa perencanaan pembelajaran adalah proses memilih, menetapkan dan mengembangkan pendekatan, metode dan teknik pembelajaran, menawarkan bahan ajar, dan menyediakan pengalaman belajar yang bermakna ser-

ta mengukur tingkat keberhasilan proses pembelajaran dalam mencapai hasil pembelajaran. Hasil penelitian secara umum menunjukkan bahwa pencapaian pelaksanaan perencanaan pembelajaran yang dilakukan oleh guru rata-rata tercapai dengan kriteria "Sangat Efektif". Hal tersebut berarti pelaksanaan perencanaan pembelajaran sudah sesuai dengan standar proses pembelajaran dimana hampir seluruh guru telah menyusun rencana pembelajaran dengan maksimal, meskipun terdapat beberapa guru yang tidak menganalisis ulang Standar Kompetensi (SK) dan Kompetensi Dasar (KD) yang ada pada Standar Isi (SI) pada penyusunan silabus atau mengkaji ulang RPP yang dibuat.

Kedua, efektivitas pelaksanaan kegiatan pembelajaran SMPN 1 Godean, SMPN 1 Kalasan dan SMPN 4 Depok secara umum telah tercapai dengan kriteria "Sangat Efektif". Pelaksanaan kegiatan pembelajaran yang dilaksanakan oleh guru telah sesuai dengan standar proses pembelajaran. Hasil penelitian menunjukkan bahwa pelaksanaan kegiatan pembelajaran yang dilakukan guru di setiap sekolah jika diurutkan skor ketercapaian pelaksanaan perencanaan pembelajaran dari yang tertinggi menurut sekolah, yaitu SMPN 4 Depok, SMPN 1 Godean, dan SMPN 1 Kalasan. Lebih lanjut, Majid (2008, p. 19) mengungkapkan bahwa proses belajar mengajar merupakan interaksi yang dilakukan antara guru dan peserta didik dalam suatu pengajaran untuk mewujudkan tujuan yang telah ditetapkan. Oleh karena itu dalam proses belajar mengajar diharapkan guru tidak mendominasi kegiatan pembelajaran yang dapat membuat peserta didik menjadi pasif. Guru berkewajiban untuk mempersiapkan metode atau strategi serta media yang sesuai dengan materi yang diajarkan dan karakteristik peserta didik sehingga dapat menarik minat peserta didik dalam mengikuti kegiatan pembelajaran.

Ketiga, efektivitas penilaian proses pembelajaran SMPN 1 Godean, SMPN 1 Kalasan dan SMPN 4 Depok yang terdiri dari tiga aspek, yaitu, pelaksanaan penilaian, pemanfaatan hasil penilaian pembelajaran

dan tindak lanjut hasil penilaian pembelajaran tercapai dengan kriteria "Sangat Efektif". Hal tersebut menunjukkan bahwa pelaksanaan penilaian pembelajaran yang dilaksanakan oleh guru telah sesuai dengan standar proses pembelajaran. Hasil penelitian menunjukkan bahwa aspek penilaian proses pembelajaran di setiap SMP dapat diurutkan skor capaian aspek penilaian proses pembelajaran mulai dari SMPN 1 Godean, SMPN 1 Kalasan dan SMPN 4 Depok. Lebih lanjut, *Permendiknas* Nomor 20 Tahun 2007 (Menteri Pendidikan Nasional Republik Indonesia, 2007b) menyatakan bahwa pelaksanaan penilaian hasil belajar peserta didik pada pendidikan dasar dan menengah dilakukan berdasarkan pada prinsip-prinsip yang meliputi: (1) sah; (2) objektif; (3) adil; (4) terpadu; (5) terbuka; (6) menyeluruh dan berkesinambungan; (7) sistematis; (8) beracuan kriteria; (9) akuntabel.

Keempat, efektivitas pengawasan atau supervisi proses pembelajaran di SMPN 1 Godean, SMPN 1 Kalasan dan SMPN 4 Depok. Prasojo & Sudiyono (2011, p. 1) mengungkapkan bahwa "supervisi merupakan kegiatan bantuan pembinaan ke arah perbaikan pembelajaran". Dari pernyataan tersebut dapat terlihat bahwa tujuan dari supervisi yaitu perbaikan pembelajaran. Hal tersebut dapat dilakukan apabila kepala sekolah melaksanakan kegiatan pengawasan secara rutin untuk mengetahui kelemahan-kelemahan proses pembelajaran yang dilakukan oleh guru.

Berdasarkan hasil penelitian pada aspek pengawasan, kepala sekolah mengisi seluruh pernyataan dari peneliti dengan jawaban "Ya", hal tersebut menunjukkan bahwa Kepala Sekolah SMPN 1 Godean telah melaksanakan pengawasan sesuai prosedur dengan baik. Namun dalam hal melakukan perencanaan dan pemetaan dalam pengawasan pembelajaran yang meliputi pemantauan, supervisi, dan evaluasi proses pembelajaran belum dilakukan. Hal tersebut menunjukkan bahwa selama ini kepala sekolah belum turut serta dalam merencanakan metode pengawasan pembelajaran dan tidak melakukan pemetaan pengawasan

pembelajaran. Pelaksanaan pengawasan pembelajaran dibagi ke dalam 4 (empat) aspek, yaitu aspek perencanaan pengawasan, pelaksanaan pengawasan, pengecekan dan tindak lanjut atau pelaporan. Dari keempat aspek tersebut aspek yang memiliki capaian terendah, yaitu aspek perencanaan pengawasan. Secara keseluruhan, pelaksanaan pengawasan di SMPN 1 Godean, SMPN 1 Kalasan dan SMPN 4 Depok memperoleh capaian dengan kriteria “Sangat Efektif”, artinya pelaksanaan pengawasan sudah sesuai dengan standar proses pembelajaran.

Selanjutnya, hasil penelitian menunjukkan bahwa terdapat delapan aspek pendukung pembelajaran di SMPN 1 Godean, SMPN 1 Kalasan, dan SMPN 4 Depok dengan skor perolehan mencapai 93,27% dengan skor 582 dari jumlah skor ideal 624, yaitu pelatihan penyusunan silabus, pelatihan penyusunan RPP, fasilitas yang memadai, lingkungan kelas yang kondusif, peserta didik yang kooperatif, ketercapaian tujuan pembelajaran, arahan kepala sekolah dalam menganalisis butir soal serta jadwal pelaksanaan evaluasi yang tepat waktu.

Hambatan pembelajaran yang dialami oleh guru diuraikan menjadi tujuh indikator, yaitu: (1) guru merasa kesulitan dalam merancang pembelajaran aktif bagi peserta didik; (2) guru merasa kesulitan dalam menentukan metode yang tepat sesuai dengan karakteristik peserta didik; (3) guru merasa kurang mampu menginovasikan media pembelajaran; (4) guru merasa kurang mampu mengaplikasikan metode yang sesuai dengan kondisi peserta didik; (5) guru merasa kesulitan untuk memotivasi kemauan belajar peserta didik. (6) guru merasa kekurangan informasi baru (misalnya materi tambahan) yang sesuai dengan usia dan tingkatan kemampuan belajar peserta didik. (7) guru merasa kesulitan mengidentifikasi kemampuan peserta didik.

Simpulan

Berdasarkan hasil analisis yang dilakukan dapat disimpulkan bahwa efektivitas implementasi standar proses pendidikan pada aspek perencanaan pembelajaran, pe-

laksanaan, penilaian, dan pengawasan di SMPN 1 Godean, SMPN 1 Kalasan dan SMPN 4 Depok telah tercapai dengan kriteria “Sangat Efektif”. Selanjutnya, diketahui bahwa terdapat delapan faktor yang mendukung terlaksananya standar proses pendidikan di sekolah, yaitu pelatihan penyusunan silabus, pelatihan penyusunan RPP, fasilitas yang memadai, lingkungan kelas yang kondusif, peserta didik yang kooperatif, ketercapaian tujuan pembelajaran, arahan kepala sekolah dalam menganalisis butir soal serta jadwal pelaksanaan evaluasi yang tepat waktu. Selain itu terdapat tujuh aspek yang menghambat pelaksanaan standar proses atau kesulitan yang dihadapi guru dalam pelaksanaan standar proses, yakni kesulitan dalam merancang pembelajaran aktif bagi peserta didik, menentukan metode yang tepat sesuai dengan karakteristik peserta didik, menginovasikan media pembelajaran, mengaplikasikan metode yang sesuai dengan kondisi peserta didik, memotivasi kemauan belajar peserta didik, memperoleh informasi baru yang sesuai dengan usia dan tingkatan kemampuan belajar peserta didik serta kesulitan dalam mengidentifikasi kemampuan peserta didik.

Berdasarkan hasil tersebut, maka disarankan bagi kepala sekolah untuk menyelenggarakan pengawasan dan evaluasi pembelajaran secara rutin, dengan terlebih dahulu melakukan perencanaan dan pemetaan, sehingga dapat meningkatkan peran guru melaksanakan standar proses pembelajaran. Kepala sekolah juga perlu menyelenggarakan pelatihan dan diskusi rutin untuk perbaikan pembelajaran agar guru dapat menyelesaikan permasalahan-permasalahan yang dihadapi dalam proses pembelajaran. Selain itu, guru diharapkan dapat mengevaluasi perencanaan pembelajaran yang disesuaikan dengan kondisi peserta didik dan karakteristik materi pelajaran secara rutin.

Daftar Pustaka

- Lestari, A. H. (2014). Rangka mutu pendidikan RI di dunia paling jeblok. *Okezone News*. Retrieved from <https://news.okezone.com/read/2014>

- /05/13/373/984246/rangking-mutu-
pendidikan-ri-di-dunia-paling-jeblok
- Majid, A. (2008). *Perencanaan pembelajaran mengembangkan standar kompetensi guru*. Bandung: PT. Remaja Rosdakarya.
- Menteri Pendidikan Nasional Republik Indonesia. Peraturan Menteri Pendidikan Nasional Republik Indonesia Nomor 41 Tahun 2007 tentang Standar Proses untuk Satuan Pendidikan Dasar dan menengah (2007).
- Menteri Pendidikan Nasional Republik Indonesia. Permendiknas nomor 20 tahun 2007 tentang standar penilaian pendidikan (2007).
- Prasojo, L. D., & Sudiyono. (2011). *Supervisi pendidikan*. Yogyakarta: Gava Media.
- Provus, M. (1969). *The discrepancy evaluation modelan approach to local program improvement and development*. Pennsylvania: PittsburgPublic School.
- Sugiyono. (2014). *Metode penelitian pendidikan pendekatan kuantitatif, kualitatif, dan R & D*. Bandung: Alfabeta.
- Usman, H. (2014). *Manajemen: teori, praktik, dan riset pendidikan* (4th ed.). Jakarta: PT Bumi Aksara.

MODEL EVALUASI IMPLEMENTASI KEBIJAKAN STANDAR PELAYANAN MINIMAL PADA SATUAN PENDIDIKAN JENJANG SD

Lilik Sabdaningtyas

FKIP-Universitas Lampung

Jl. Prof. Sumantri Brojonegoro No.1, Rajabasa, Gedong Meneng, Rajabasa, Kota Bandar
Lampung, Lampung 35145, Indonesia

Email: lilik_sabdaningtyas@yahoo.co.id

Abstrak

Tujuan penelitian ini untuk menghasilkan model evaluasi implementasi kebijakan standar pelayanan minimal (SPM) jenjang sekolah dasar (SD). Hasil evaluasi diharapkan dapat memberikan rekomendasi kepada Dinas pendidikan dan para praktisi pendidikan. Desain penelitian yang digunakan adalah *non experimental*. Desain penelitian ini terdiri dari 5 variabel *laten* dan 33 variabel *manifest*. Populasinya mencakup 307 SD di 11 Kecamatan wilayah Kabupaten Pesawaran. Sampel sebanyak 24 SD pada 4 Kecamatan diambil secara purposive. Instrumen pengumpulan data yang digunakan adalah kuesioner yang divalidasi dan *content validity* melalui *Focus Group Discussion* (FGD). Data yang terkumpul dianalisis menggunakan *Structural Equation Models* (SEM). Hasil penelitian menunjukkan bahwa Kondisi Lingkungan Sekolah (KLS), Pelaksana Kebijakan Sekolah (PKS), Kelompok Sasaran Kebijakan (KSK), dan Proses Pengelolaan Organisasi Sekolah (PPOS) merupakan faktor dinamis pembentuk model yang tepat untuk mengevaluasi IP-SPM SD.

Kata kunci: *model evaluasi, implementasi kebijakan, IP-SPM-SD*

EVALUATION MODEL OF MINIMUM SERVICE STANDARD POLICY IMPLEMENTATION AT ELEMENTARY EDUCATION LEVEL

Abstract

The purpose of this research is to produce evaluation model of minimum service standard implementation at elementary education level. The evaluation results are expected to provide recommendations to education authorities and practitioners. The research employed non experimental design. This research design consisted of five latent variables and 33 manifest variables. Its population includes 307 elementary schools in 11 sub-districts of Pesawaran Regency. Purposive sampling technique was used to select 24 elementary schools in four sub-districts. The data collection instrument used in this study was questionnaire which was validated by content validity through focus group discussion (FGD). The collected data were analyzed using structural equation model (SEM). The results of the research show that the condition of school environment, school policy executors, policy-target group, and school organizational management process are the appropriate dynamic modeling factor to evaluate the achievement index of minimum service standard at elementary education level. Hasil penelitian menunjukkan bahwa Kondisi Lingkungan Sekolah (KLS), Pelaksana Kebijakan Sekolah (PKS), Kelompok Sasaran Kebijakan (KSK), dan Proses Pengelolaan Organisasi Sekolah (PPOS) merupakan faktor dinamis pembentuk model yang tepat untuk mengevaluasi IP-SPM SD

Keywords: *evaluation model, policy implementation, IP-SPM-S*

Permalink/DOI: <http://dx.doi.org/10.21831/pep.v22i1.19638>

Pendahuluan

Reformasi dalam penyelenggaraan pemerintahan di Indonesia telah menyebabkan terjadinya sejumlah perubahan penting dan mendasar dalam tata kelola pemerintahan yang pada akhirnya berimplikasi pada penyelenggaraan pelayanan publik di daerah. Lembaga pendidikan yang merupakan lembaga pelayanan publik juga mengalami perubahan.

Terkait dengan penyelenggaraan pelayanan publik di bidang pendidikan, pemerintah telah memformulasikan Peraturan Menteri Pendidikan dan Kebudayaan (Permendikbud) No 23 Tahun 2013 tentang Standar Pelayanan Minimal (SPM) Pendidikan Dasar di Kabupaten/Kota. Permendikbud tersebut telah diformulasikan sejak tahun 2011 dan ditargetkan tahun 2014 seluruh pelayanan minimal pendidikan Dasar di seluruh Kabupaten/Kota telah terpenuhi (Menteri Pendidikan dan Kebudayaan Republik Indonesia, 2013).

Hingga saat ini yakni tahun 2018 implementasi kebijakan tersebut telah berjalan selama kurun waktu tujuh tahun. Untuk mengetahui efektivitas implementasinya maka perlu dilakukan evaluasi, dengan harapan hasil evaluasi bisa dijadikan pemetaan dalam rangka Pemerintah Daerah mendistribusikan bantuan kepada sekolah-sekolah.

Mengacu pada Permendikbud No 23 Tahun 2013, SPM pendidikan dasar adalah merupakan tolok ukur kinerja pelayanan pendidikan dasar melalui jalur pendidikan formal yang diselenggarakan oleh Pemerintah Kabupaten/Kota. SPM pendidikan merupakan ketentuan tentang jumlah dan mutu layanan pendidikan yang diselenggarakan oleh Pemerintah Kabupaten/Kota, Kantor Wilayah Kementerian Agama, dan Kantor Kementerian Agama Kabupaten/Kota secara langsung maupun secara tidak langsung melalui sekolah dan madrasah. Tujuan penerapan SPM dimaksudkan untuk memastikan bahwa di setiap sekolah dan madrasah terpenuhi kondisi minimum yang dibutuhkan untuk menjamin terselenggaranya proses pembelajaran yang memadai. SPM diharapkan mampu mempersempit kesenjangan mutu pendidikan yang kedepannya juga

diharapkan berimplikasi pada mengecilnya kesenjangan sosial ekonomi.

Menyikapi tentang implementasi kebijakan pemerintah yang tertuang dalam Permendikbud No 23 Tahun 2013 tentang SPM Pendidikan Dasar di Kabupaten/Kota. Pemerintah telah melakukan berbagai upaya agar target bisa tercapai. Beberapa upaya tersebut diantaranya adalah rehabilitasi sarana dan prasarana pada satuan sekolah, pelatihan guru, sertifikasi guru dan lain-lain. Namun kenyataan di lapangan ternyata belum menggembirakan, hal ini terlihat masih adanya beberapa permasalahan sehubungan dengan implementasi kebijakan SPM.

Beberapa permasalahan tersebut diantaranya adalah; (1) implementasi kebijakan tersebut telah memasuki tahun ke tujuh dan telah melebihi target yang ditetapkan Pemerintah yakni tahun 2014 SPM harus tercapai. Ini bisa dimaknakan bahwa target pencapaiannya tidak terpenuhi, seharusnya pada tahun 2014 semua jenjang SD telah memenuhi standar pelayanan dasar; (2) pelaksanaan SPM tentu tidaklah semudah membalikkan tangan. Hal ini mengingat bahwa tidak menjadi jaminan bagi setiap daerah dan satuan pendidikan mampu memberikan pelayanan yang berkualitas kepada masyarakat dalam waktu yang relatif singkat yang sesuai target pemerintah yakni 3 tahun berjalan; (3) disisi lain kesiapan daerah maupun satuan sekolah untuk melaksanakan tugas pelayanan ini relatif terbatas baik dari segi sumber daya, kelembagaan, maupun keuangannya; (4) permasalahan lain adalah disparitas kapasitas daerah maupun sekolah dalam melaksanakan pelayanan dasar di era otonomi daerah menunjukkan tingkat kesenjangan yang cukup tinggi, namun daerah dan satuan sekolah harus tetap melaksanakan SPM yang sama secara nasional; (5) fakta lain, sebagaimana hasil penelitian yang dilakukan oleh Adawiah (2011) menunjukkan bahwa SPM pendidikan dasar di Kabupaten Balangan umumnya masih jauh dari harapan. Solongi (2015) yang mengacu pada data hasil pemetaan Kemendikbud menunjukkan bahwa dari 148.061 SD/MI/SMP/MTS yang tersebar di seluruh tanah air yang dihuni

sebanyak 26.119.000 murid, masih ada 75% yang tidak memenuhi SPM. Terdapat puluhan ribu sekolah yang masih dalam kondisi rusak. Jumlah ruang kelas SD dan SMP yang rusak berat diperkirakan mencapai 739,741. Kondisi ini mengindikasikan bahwa SPM belum sepenuhnya terimplementasi di seluruh daerah. Hasil penelitian Hermawan (2011, pp. 619–634), menunjukkan bahwa hanya 23 sekolah (52,27%) yang telah mencapai standar nasional pendidikan (SNP), sedangkan 21 sekolah (47,73%) lainnya belum mencapai. Berdasarkan fakta empirik ini, perlu dilakukan evaluasi untuk mengetahui faktor-faktor apa yang menjadi penyebab sehingga target SPM belum terpenuhi.

Permasalahan tersebut sangat menarik perhatian untuk diteliti mengingat: (1) kebijakan tentang implementasi SPM khususnya pada satuan sekolah jenjang SD merupakan masalah yang aktual, (2) kebijakan dan program yang tidak dievaluasi implementasinya maka tidak akan diketahui kelemahan dan kelebihan dalam rangka pengembangan, kelanjutan, atau pemberhentian program tersebut.

Berdasarkan permasalahan dan argumen tersebut perlu dikembangkan sebuah model evaluasi yang tepat yakni; “Model Evaluasi Implementasi Kebijakan SPM pada Satuan Pendidikan Jenjang SD di Kabupaten Pesawaran”.

Menurut Dunn (1998, pp. 608–610) secara umum istilah evaluasi dapat disamakan dengan penaksiran (*appraisal*), pemberian angka (*rating*) dan penilaian (*assessment*) sebagai usaha untuk menganalisis hasil kebijakan. Rossi & Freeman (1982, p. 4) mengklasifikasikan tiga besar evaluasi kebijakan yaitu: (a) evaluasi pada tahap rancangan dan konseptualisasi program; (b) evaluasi pada tahap implementasi program (*monitoring* dan akuntabilitas); (c) evaluasi pada tahap kegunaan program (pengukuran *efisiensi* dan *impact*). Kemudian berdasar fungsinya, evaluasi kebijakan dibedakan menjadi empat. Pertama, Eksplanasi. Dalam eksplanasi dapat dibuat suatu gambaran tentang pola-pola hubungan antarberbagai dimensi realita yang dapat diamatinya. Melalui Eksplanasi dapat

diketahui variabel-variabel kebijakan yang mempengaruhi keberhasilan suatu kebijakan. Dengan demikian dapat diidentifikasi tentang tujuan-tujuan dari program utama yang potensial untuk tercapai, mengapa tujuan itu harus dicapai dan bagaimana mencapainya. Kedua, Kepatuhan. Kepatuhan berfungsi untuk melihat apakah tindakan yang dilakukan oleh pelaku (birokrasi maupun pelaku lain) sesuai dengan standar dan prosedur yang ditetapkan oleh kebijakan. Ketiga, Fungsi *auditing*. Ini berfungsi untuk melihat apakah *output* benar-benar sampai ke tangan kelompok sasaran maupun penerima lain (individu, organisasi, dll) yang dimaksud oleh pembuat kebijakan. Keempat, *Akunting*. Fungsi ini adalah untuk melihat apakah akibat sosial ekonomi dari kebijakan tersebut. Lebih lanjut dikemukakan bahwa tujuan evaluasi kebijakan pada dasarnya adalah: (1) untuk mengetahui proses pembuatan kebijakan, (2) untuk mengetahui proses implementasi, (3) untuk mengetahui konsekuensi kebijakan, dan (4) untuk mengetahui efektivitas dampak kebijakan (Wibawa, Prabuningrat, & Pramusinto, 1994, pp. 9–10).

Untuk melakukan evaluasi diperlukan model evaluasi tertentu. Worthen & Sanders (1973, p. 20) mengemukakan evaluasi program adalah suatu proses mengidentifikasi dan mengumpulkan informasi untuk membantu para pengambil keputusan dalam memilih berbagai alternatif keputusan. Sudjana (2006, p. 51) mengemukakan, model evaluasi program mencakup lebih dari 50 jenis yang telah dan sedang digunakan dalam evaluasi program. Model-model tersebut dapat dikelompokkan ke dalam enam kategori yaitu: (1) model evaluasi terfokus pada pengambilan keputusan; (2) model evaluasi terhadap unsur-unsur program; (3) model evaluasi terhadap jenis/tipe kegiatan program; (4) model evaluasi terhadap proses pelaksanaan program; (5) model evaluasi terhadap pencapaian tujuan program; dan (6) model evaluasi terhadap hasil dan pengaruh program. Berdasarkan telaah tentang evaluasi kebijakan (*policy evaluation*) di atas, maka harus difahami bahwa untuk me-

lakukan evaluasi kebijakan kita harus melakukan pilihan-pilihan terhadap beberapa hal agar pelaksanaan evaluasi terfokus dan sesuai dengan fenomena yang akan kita evaluasi. Pilihan-pilihan yang dimaksud adalah pilihan evaluasi menurut tahapannya, menurut bentuk, menurut fungsi, menurut tujuannya, menurut model, dan menurut *scope*-nya.

Istilah kebijakan secara sederhana dapat diartikan sebagai “pedoman untuk bertindak”. Kebijakan dalam maknanya seperti ini bisa berupa suatu deklarasi mengenai suatu dasar pedoman bertindak, suatu arah tindakan tertentu, suatu program mengenai aktivitas-aktivitas tertentu atau suatu rencana. Dengan kata lain kebijakan juga dapat merupakan ketentuan-ketentuan yang harus dijadikan acuan, pedoman, pegangan, dan petunjuk serta cara bagi setiap usaha dan kegiatan sehingga tercapai kelancaran dan keterpaduan dalam mencapai tujuan tertentu. Implementasi kebijakan merupakan salah satu tahapan dalam keseluruhan proses kebijakan yaitu tahap formulasi, tahap implementasi, dan tahap evaluasi yang berlangsung dalam suatu sistem kebijakan yang kompleks dan dinamik serta akan menentukan berhasil atau gagalnya suatu kebijakan. Oleh karena itu implementasi kebijakan adalah merupakan suatu tahapan penting dalam proses kebijakan untuk dikaji agar bisa mengantisipasi terhadap keberhasilan kebijakan itu sendiri. Definisi sederhana dari istilah implementasi adalah “penerapan atau pelaksanaan”. Implementasi kebijakan bersifat interaktif dengan kegiatan kebijakan yang mendahuluinya yakni formulasi kebijakan. Oleh karena itu implementasi kebijakan dapat dipandang sebagai sebuah proses interaksi antara suatu perangkat tujuan dan tindakan yang harus dilakukan.

Ada tiga kegiatan penting yang berkaitan dengan tahapan implementasi kebijakan yaitu; organisasi, interpretasi, dan penerapan. Dalam hal ini organisasi adalah memerlukan perhatian yang paling utama karena sangat penting bagi pembahasan konsep birokrasi serta bentuk-bentuknya yang terkait sekali dengan sebuah kebijakan (Jones, 1994, pp. 294–296). Organisasi yang

dimaksudkan oleh Jones adalah pembentukan atau penataan kembali sumber daya, unit-unit serta metode untuk menjadikan program berjalan. Dalam pelaksanaan sebuah kebijakan tentu diperlukan adanya organisasi agar pekerjaan dapat dilaksanakan.

Selain hal tersebut bahwa implementasi sebuah kebijakan akan sangat dipengaruhi berbagai faktor. Faktor-faktor yang berpengaruh terhadap implementasi kebijakan yang dikemukakan beberapa ahli diantaranya; Grindle (1980), Mazmanian & Sabatier (1986), Edwards III (1980), Van Meter & Van Horn (1973), Cheema & Rondinelli (1983), dan Koster (2000). Menurut peneliti, beberapa faktor yang mempengaruhi implementasi kebijakan sebagaimana pendapat para ahli di atas, jika dikelompokkan terdapat kesamaan 4 (empat) aspek utama sebagai faktor pendukung implementasi kebijakan yaitu: (a) aspek yang terkait dengan kondisi lingkungan; (b) aspek yang terkait dengan isi kebijakan; (c) aspek yang terkait dengan karakteristik organisasi pelaksana; (d) aspek yang terkait dengan karakteristik kelompok sasaran (*target group*). Disamping empat aspek tersebut, terdapat satu aspek penting lain yang merupakan pendukung implementasi kebijakan yakni proses pengelolaan organisasi (Sabdaningtyas, 2010). Dengan mengacu pada beberapa pendapat tersebut di atas, bahwa kondisi lingkungan, pelaksana kebijakan, kelompok sasaran kebijakan, dan proses pengelolaan organisasi merupakan faktor-faktor pendukung implementasi kebijakan SPM pada satuan pendidikan jenjang SD.

Indikator Model Evaluasi Implementasi Kebijakan SPM SD

Indikator adalah karakteristik yang dapat diobservasi secara langsung sebagai ganti dari karakteristik yang tak dapat diamati secara langsung dan digunakan sebagai definisi operasional atas suatu variabel (Dunn, 1998, p. 597). Secara sistemik, model evaluasi implementasi kebijakan bukan merupakan fenomena yang berdiri sendiri melainkan merupakan hasil transformasi berbagai *inputs* melalui suatu komponen *throughput*

(proses) tertentu, maka evaluasi implementasi kebijakan juga dapat menyentuh berbagai indikator baik pada komponen *input* maupun komponen *throughput*, dan dipengaruhi pula oleh maksud dan tujuan evaluasi itu dilakukan yakni bertujuan untuk pemantauan, pengawasan atau pertanggung jawaban. Dalam penelitian ini akan dilakukan evaluasi implementasi kebijakan SPM pada satuan pendidikan jenjang SD yang termasuk kegiatan yang bertujuan untuk pemantauan.

Mustopadidjaja (2002, p. 45) menguraikan evaluasi implementasi kebijakan yang dilakukan dengan tujuan pemantauan dimaksudkan untuk mendapatkan informasi dini mengenai perkembangan pelaksanaan pada momentum atau dalam jangka waktu tertentu sehingga dapat diketahui hal-hal yang perlu diperbaiki mengenai sistem dan proses pelaksanaan kebijakan tersebut agar pelaksanaan kebijakan dapat berjalan optimal. Disamping memuat tentang gambaran perkembangan pelaksanaan, dalam evaluasi pemantauan juga memuat identifikasi kelemahan-kelemahan, penyimpangan yang terjadi serta potensi atau daya dukung yang ada selama proses pelaksanaan kebijakan tersebut. Berdasarkan telaah beberapa teori para ahli evaluasi implementasi kebijakan, maka indikator yang digunakan untuk mengukur tiap-tiap variabel dalam penelitian ini menggunakan pola pikir sistemik yakni dengan mengangkat beberapa komponen yang meliputi *Policy input*, *Policy cycle*, dan *Policy output*. peneliti menyingkat dan menamakan model ini sebagai model IPO (Input, Proses, dan Output). Komponen IPO dalam penelitian ini diuraikan sebagai berikut.

Pertama, komponen *Policy Output* (keluaran kebijakan) adalah barang, jasa atau sumber daya yang diterima oleh kelompok sasaran dan pihak yang menerima akibat (Dunn, 1998, p. 597). Variabel dalam komponen *policy output* dalam penelitian ini adalah IP (Indek Pencapaian) SPM SD. Indikator yang digunakan untuk mengevaluasi komponen ini mengacu pada Lampiran II Permendikbud No 23 tahun 2013 tentang SPM Pendidikan Dasar di Kabupaten/ Kota.

Kedua, komponen *Policy Cycle* (Proses). Proses merupakan rangkaian kegiatan yang meliputi tiga kelompok kegiatan utama yaitu formulasi kebijakan, implementasi kebijakan, dan evaluasi kinerja kebijakan (Mustopadidjaja, 2002, p. 3). *Policy cycle* dalam penelitian ini hanya difokuskan pada tahap implementasi kebijakan. Dalam bukunya, Dunn (1998, p. 598) mengartikan implementasi sebagai kegiatan dan sikap administratif, organisasional yang menentukan transformasi masukan kebijakan menjadi keluaran. Dalam penelitian ini implementasi diartikan sebagai suatu proses pengelolaan input sehingga variabel yang dipilih dalam komponen *policy cycle* adalah variabel proses pengelolaan organisasi sekolah (PPOS). Variabel PPOS yang dimaksud adalah proses pengelolaan organisasi sekolah oleh birokrat sekolah dalam rangka melaksanakan IP-SPM SD. Komponen proses ini pada dasarnya mempertanyakan apakah proses pengelolaan organisasi sekolah telah dilaksanakan sesuai dengan ketentuan yang berlaku. Proses pengelolaan organisasi sekolah efektif bila mempertimbangkan tiga komponen yakni: (1) struktur organisasi jelas, (2) pola hubungan yang dijalin dalam organisasi tersebut baik, dan (3) ketepatan jenis organisasi yang diterapkan. Pola hubungan yang dijalin dalam organisasi sekolah tidak lagi bersifat hierarki tetapi lebih bersifat fungsional, konsultatif, kemitraan dan koordinatif, dengan demikian perlu dikembangkan jenis organisasi yang non struktural, fungsional dan organik semacam *corporate organization* (Satori, 1999, p. 17). Berdasarkan uraian di atas, maka indikator yang digunakan dalam variabel PPOS adalah: (1) Struktur organisasi yang meliputi kompleksitas, formalitas, dan sentralisasi; (2) Pola hubungan yang diciptakan dalam organisasi sekolah hendaknya fungsional, konsultatif, kemitraan, dan koordinatif jadi bukan lagi pola hubungan yang otoriter; (3) Jenis organisasi yang digunakan hendaknya organisasi yang demokratis atau non struktural dan fungsional yang didasarkan atas profesi keahliannya.

Ketiga, Komponen *Policy Input*. *Policy input* adalah beberapa faktor dinamik yang

saling berinteraksi dan harus tersedia karena dibutuhkan untuk berlangsungnya proses kebijakan (Mustopadidjaja, 2002, p. 3). Faktor-faktor dinamik yang harus tersedia sebagai komponen *policy input* dalam penelitian ini meliputi: variabel kondisi lingkungan, variabel pelaksana, dan variabel kelompok sasaran. Komponen *policy input* ini pada dasarnya adalah mempertanyakan apakah lingkungan kebijakan, pelaksana, dan kelompok sasaran sebagai input kebijakan mendukung terhadap kinerja implementasi kebijakan SPM pada satuan pendidikan. Masing-masing variabel dalam komponen *policy input* ini akan diukur dengan beberapa indikator yang diuraikan sebagai berikut.

Pertama, variabel Kondisi Lingkungan Sekolah (KLS). Variabel kondisi lingkungan sekolah diukur dengan tiga indikator yakni: (a) Kondisi geografis, ini akan dilihat letak/keberadaan satuan sekolah secara geografis; (b) Kondisi sosial dan kondisi ekonomi, ini akan dilihat dari kondisi rata-rata tingkat pendidikan dan kondisi rata-rata tingkat ekonomi masyarakat/orang tua siswa pada satuan sekolah; (c) Kondisi politik, ini akan dilihat besar kecilnya dukungan para elit politik dalam ikut serta mewujudkan terlaksananya SPM sekolah.

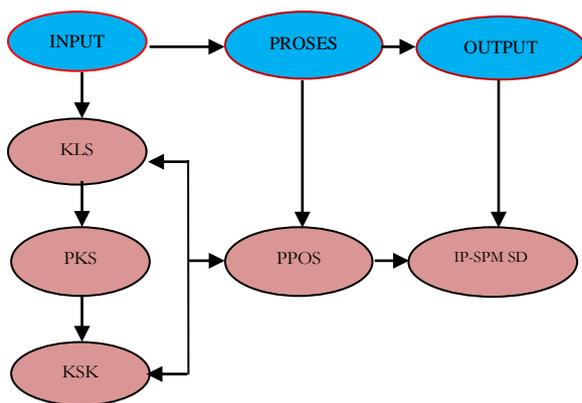
Kedua, variabel Pelaksana Kebijakan Sekolah (PKS). Variabel pelaksana kebijakan sekolah mengukur variabel tersebut adalah: (a) Aspek *communication*. Dengan indikator bahwa antar pelaksana kebijakan pada satuan pendidikan tidak terjadi perbedaan pandangan terhadap kebijakan program SPM pendidikan pada satuan sekolah yang akan dilaksanakan, yang disebabkan karena komunikasi antar pelaksana kebijakan lancar dan mudah dimengerti semua pihak; (b) Aspek *resources*. Dengan indikator ketaatan/kepatuhan personil-personil pelaksana dalam menjalankan tugasnya sesuai dengan kedudukan, tugas dan fungsinya sebagai pelaksana program; (c) Aspek sikap pelaksana. Dengan indikator para pelaksana bersikap positif terhadap kebijakan yang akan dilaksanakan, ini ditunjukkan dengan adanya para pelaksana di tingkat satuan sekolah mampu

nyai persepsi yang baik dan sama terhadap implementasi kebijakan SPM sekolah.

Ketiga, variabel Kelompok Sasaran Kebijakan (KSK) Kelompok sasaran kebijakan adalah orang atau sekelompok orang atau organisasi-organisasi dalam masyarakat yang perilaku dan atau keadaannya akan dipengaruhi oleh kebijakan (Mustopadidjaja, 2002, p. 8). Variabel ini diukur melalui tiga indikator yaitu: (a) Aspek *targetgroup* menerima kebijakan yang diimplementasikan, yang indikatornya adalah mereka merasa membutuhkan dan juga memperoleh manfaat dari kebijakan yang diimplementasikan; (b) Aspek *targetgroup* bersikap netral/acuh tak acuh, dengan indikator mereka tidak menyadari manfaat dari kebijakan SPM. c) Aspek *targetgroup* menolak kebijakan yang diimplementasikan, dengan indikator semata-mata karena mereka tidak menyukai perubahan karena sama sekali tidak memperoleh manfaat dari kebijakan yang diimplementasikan.

Kerangka Pikir Penelitian

Sebuah kebijakan akan menjadi impian yang tersimpan sebagai arsip belaka bila tidak diimplementasikan dengan efektif. Implementasi kebijakan merupakan aspek penting yang bisa menentukan berhasil atau gagalnya suatu kebijakan. Kebijakan tentang SPM SD telah diformulasikan sejak tahun 2011 dan hingga saat ini implementasinya telah berjalan selama kurun waktu lima (5) tahun. Untuk mengetahui efektivitas implementasi kebijakan SPM pada satuan pendidikan yang saat ini sedang berlangsung maka perlu dilakukan evaluasi. Untuk keperluan evaluasi diperlukan model evaluasi yang tepat agar hasil evaluasi menjadi akurat. Oleh karena itu perlu dikembangkan model evaluasi yang tepat untuk mengukur Indeks Pencapaian (IP) implementasi kebijakan SPM pada satuan pendidikan jenjang SD.



Gambar 1. Kerangka Pikir Penelitian

Pengembangan model implementasi kebijakan SPM SD yang dimaksud adalah membuat model yang didasarkan atas pengkajian terhadap kerangka teoritis yang kemudian diuji berdasarkan fakta empiris. Dalam konsep teoritis, untuk mencapai tujuan kebijakan (*policy output*) harus dilakukan aksi atau tindakan yang berupa penghimpunan sumber daya yang disebut input kebijakan (*policy input*) dan pengelolaan sumber daya yang disebut siklus kebijakan (*policy cycle/ Proses*) yang merupakan tahapan formulasi, implementasi, dan evaluasi kebijakan. Penelitian ini hanya memfokuskan pada tahapan implementasi kebijakan (*policy implementation*). Selanjutnya bahwa untuk menghasilkan implementasi kebijakan yang efektif maka diperlukan faktor-faktor dinamik yang terdiri dari: kondisi lingkungan kebijakan, pelaksana kebijakan, kelompok sasaran kebijakan, yang ketiganya merupakan input kebijakan dan akan berperan dalam penyelenggaraan siklus/proses kebijakan (*policy cycle*). Berdasarkan konsep teoritis tersebut, maka kerangka pikir penelitian ini dapat dilihat dalam bagan sebagaimana Gambar 1.

Hipotesis Penelitian

Beberapa hipotesis penelitian diuraikan sebagai berikut. Pertama, kondisi lingkungan sekolah, pelaksana kebijakan sekolah, kelompok sasaran kebijakan, dan proses pengelolaan organisasi sekolah merupakan faktor-faktor dinamis yang membentuk model yang tepat untuk mengevaluasi indeks pencapaian standar pelayanan minimal (IP-SPM) satuan pendidikan jenjang SD

Kedua, proses pengelolaan organisasi sekolah merupakan pendukung langsung yang signifikan terhadap indeks pencapaian standar pelayanan minimal pada satuan pendidikan jenjang SD. Ketiga, proses pengelolaan organisasi sekolah didukung langsung secara signifikan oleh kondisi lingkungan sekolah, pelaksana kebijakan sekolah, dan kelompok sasaran kebijakan sekolah.

Keempat, indeks pencapaian standar pelayanan minimal pada satuan pendidikan jenjang SD didukung langsung secara signifikan oleh kondisi lingkungan sekolah, pelaksana kebijakan, kelompok sasaran kebijakan, dan proses pengelolaan organisasi sekolah. Kelima, kondisi lingkungan sekolah, pelaksana kebijakan, dan kelompok sasaran kebijakan merupakan pendukung tidak langsung secara signifikan terhadap indeks pencapaian standar pelayanan minimal pada satuan pendidikan jenjang SD melalui proses pengelolaan organisasi sekolah. Berdasarkan uraian yang telah disampaikan tersebut, maka penelitian ini bertujuan untuk menghasilkan model evaluasi implementasi kebijakan SPM jenjang SD.

Metode Penelitian

Desain penelitian yang dirancang dalam penelitian ini merupakan desain penelitian *expost facto* atau *desain non experimental*. Sebagaimana pendapat Isaac & Michael (1981), penelitian *expost facto* mencoba menentukan suatu sebab dari sesuatu yang sudah terjadi, oleh karena itu desain penelitian yang demikian tidak ada perlakuan/*treatment* yang dilakukan peneliti.

Variabel dalam penelitian dijabarkan berdasar tiga komponen yakni komponen *Input*, *Proses* dan *output* yang masing-masing komponen terdiri dari variable *laten* dan masing-masing variabel *laten* terdiri dari beberapa variabel *manifest*. Sehingga dalam penelitian ini terdapat 5 variabel *laten* dan 33 variabel *manifest*. Lima variabel laten terdiri dari variabel eksogen dan variabel endogen. Variabel eksogen meliputi (1) KLS; (2) PKS; (3) KSK. Variabel endogen meliputi: (1) PPOS, dan (2) IP-SPM-SD.

Unit analisis penelitian ini adalah satuan pendidikan jenjang SD, sedangkan populasinya seluruh SD di Kabupaten Pesawaran yakni sebanyak 307 SD yang tersebar pada 11 Kecamatan. Pengambilan sampel dalam penelitian ini dilakukan dengan teknik purposif *sampling* yakni sebanyak 24 SD pada 4 kecamatan. Responden penelitiannya para guru dan kepala sekolah, sebagai *target group* SPM dan sekaligus sebagai *implementor* kebijakan SPM. Untuk setiap sekolah sampel akan diambil satu orang kepala sekolah, dan 2 orang guru yang terdiri dari 1 orang guru senior dan 1 orang lagi guru junior. Teknik pengumpulan data utama menggunakan kuesioner, sedangkan wawancara, dan data dokumen akan digunakan sebagai teknik pelengkap.

Kriteria pengujian model dalam penelitian ini digunakan pendekatan SEM dengan program LISREL. LISREL terdiri dari dua kelompok persamaan/model yakni; model struktural (*structural model*) dan model pengukuran (*measurement model*). Evaluasi kesesuaian model (*model fit*). Digunakan kriteria *overall fit* yakni menguji; (a) *Chi Square dan Probabilitas*, dan (b) *Root Mean Square Error of Approximation (RMSEA)*.

Prosedur yang dilakukan untuk mengembangkan model dalam penelitian ini mengikuti langkah-langkah pemikiran dalam pemodelan SEM. Muhadjir (2007, pp. 358–359) menguraikan; SEM menggunakan pemodelan struktural linier dan penstrukturannya menggunakan persamaan dalam matematika. Pemodelan struktural ini dikembangkan dari tuntutan ilmu mutakhir bahwa kebenaran teoritik dituntut terstruktur (*structured*), dan kebenaran empirik *eksperimental* dituntut struktural pula. Dengan SEM dimaksudkan agar dapat membangun dan mengembangkan kebenaran konstruk (*constructed truth*).

Salah satu pemodelan struktural adalah LISREL. LISREL merupakan model matematik formal yang *substantive content* dalam setiap aplikasinya. Bentuk umum model LISREL terdiri atas seperangkat persamaan linier struktural. Variabel-variabelnya dalam sistem persamaan dengan mengangkat

observed variables dan *latent variables*. Konstruk modelnya berasumsi bahwa ada struktur kausal diantara sejumlah *latent variable*, dan *observed variables* dijadikan indikator atau *simpton* dari *latent variable* tertentu. Dengan mengacu pada langkah-langkah pembuatan pemodelan menurut Ferdinand (2000, p. 22), Stoelting (Narimawati & Sarwono, 2007, p. 39), dan Ghozali & Fuad (2005, p. 9), maka langkah pengembangan model dalam penelitian ini sebagaimana terurai berikut.

Pertama, pengembangan teori/konseptualisasi model. Teori yang digunakan untuk mengembangkan model dalam penelitian ini menggunakan teori berfikir sistem yang komponennya terdiri dari: *policy input*, *policy cycle*, dan *policy output* yang oleh peneliti disebut sebagai model ICO (*Input, Cycle, dan Output*). Atau disebut juga dengan model IPO (*Input, Proses, dan Output*). Kedua, menjabarkan Teori ke dalam Komponen, Variabel dan Indikator. Pada langkah ke dua ini peneliti membuat kisi-kisi instrumen sebagai dasar untuk pembuatan draf instrumen penelitian.

Ketiga, menyusun instrumen dan validasi. Langkah awal dalam kegiatan ini adalah membuat draf instrumen yang kemudian divalidasikan dengan melakukan *Focus Group Discusion* (FGD) bersama para praktisi pendidikan yakni Kepala Unit Pelaksana Teknis Pendidikan (KUPTP), kepala sekolah, komite sekolah, dan para guru di wilayah dinas pendidikan Kabupaten Pesawaran. Partisipan FGD sebanyak 30 orang dan hasil FGD menunjukkan bahwa sebanyak > 50% partisipan FGD menyatakan setuju dan sangat setuju dengan draf instrumen yang dirancang. Keempat, pengembangan diagram alur dan desain model. Pada langkah ini dikembangkan desain model dengan pola berfikir sistem.

Kelima, konversi Diagram Alur ke dalam Persamaan Struktural. Atas dasar disain model yang telah dirancang, bahwa variabel PPOS memiliki fungsi ganda, dimana selain berperan sebagai variabel eksogen terhadap IP-SPM juga berperan sebagai variabel endogen terhadap KLS, PKS, dan KSK. Jadi secara struktural ada tiga model yang

dijabarkan dalam penelitian ini. Spesifikasi model struktural berturut-turut adalah; (a) model struktural dukungan PPOS terhadap IP-SPM, (b) Model struktural dukungan KLS, PKS, dan KSK terhadap PPOS, (c) Model struktural dukungan KLS, PKS, dan KSK terhadap IP-SPM.

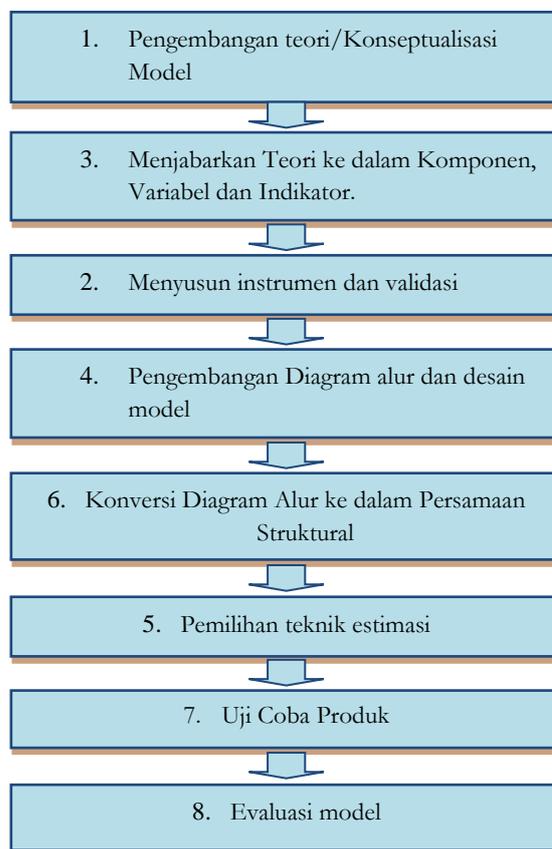
Keenam, pemilihan teknik estimasi. Tahap ini merupakan pemilihan terhadap nilai parameter awal yang bebas dipilih untuk menentukan matrik kovarian populasi yang diestimasi dari model tersebut. Terdapat beberapa metode estimasi yang dapat digunakan dalam SEM, diantaranya menurut Joreskog & Sarbon (1996, p. 17), Narimawati & Sarwono (2007, p. 44), Muhadjir (2007, pp. 359–360). Mengingat banyaknya metode estimasi yang dapat digunakan dalam SEM, dengan mengutip pendapat Hair, Tatham, Anderson, & Black (1998, p. 605) bahwa dengan metode estimasi *Maximum Likelihood* (ML) adalah merupakan metode yang paling umum digunakan dalam SEM, dinyatakan bahwa ukuran sampel 50 sudah cukup layak. Atas dasar pendapat Hair maka peneliti memilih metode estimasi *Maximum Likelihood* (ML).

Ketujuh, Uji Coba Produk. Produk yang dihasilkan dalam penelitian ini adalah: “Model Evaluasi Kebijakan Standar Pelayanan Minimal Satuan Pendidikan Jenjang SD”. Agar mendapatkan model yang tepat, maka harus dilakukan pengujian kecocokan antara model desain secara teoritis yang diajukan dengan data empiris. Oleh karena itu uji coba produk dilakukan dengan menguji kecocokan model teoritik dengan fakta empirik di lapangan. Pengujian kecocokan model dikenakan pada 24 SD yang tersebar pada 4 Kecamatan di wilayah Kabupaten Pesawaran dengan responden sebanyak 72 orang yang terdiri dari 24 orang kepala sekolah, 24 orang guru senior dan 24 orang guru junior.

Kedelapan, Evaluasi model. Evaluasi dilakukan dengan menggunakan kriteria keselarasan (*goodness of fit*). Uji keselarasan model akan dilakukan dalam tiga tahapan yakni; (1) uji kesesuaian model pengukuran, (2) uji kesesuaian model struktural, dan (3) uji mo-

del secara keseluruhan. Kriteria keselarasan yang digunakan dalam penelitian ini adalah; (1) *Chi-Square* (X^2), dan *Probabilitas* (p); (2). *Root Mean Square Error of Approximation* (RMSEA).

Secara ringkas langkah pengembangan model tertuang dalam Gambar 2.



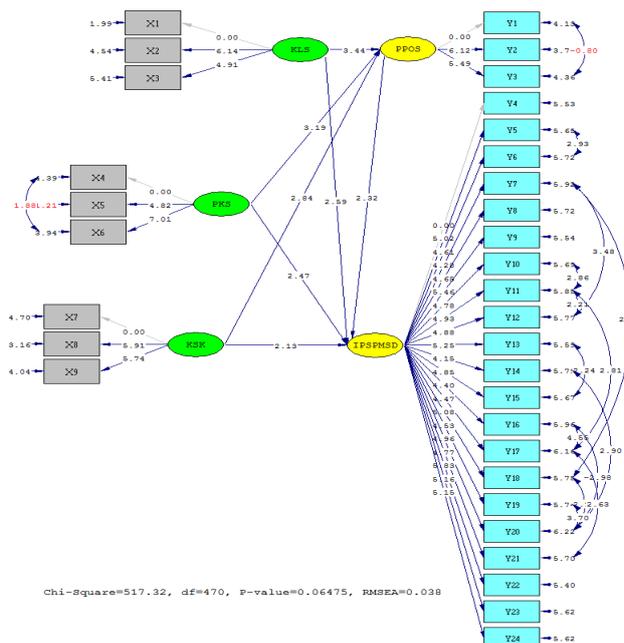
Gambar 2. Langkah Pengembangan Model

Hasil Penelitian

Uji Model Keseluruhan (*Overall/Fit*)

Hasil pengujian *Goodness of Fit Index* nampak bahwa, *Chi-Square* (X^2) dan *Probability* (p) sebesar 517.32 dengan p 0.06457 > p 0.05. Kemudian *Root Mean Square Error of Approximation* (RMSEA) sebesar $0,038 \leq p$ 0.05. Hasil ini menunjukkan bahwa model terpenuhi yang berarti bahwa model teoritik yang dihipotesiskan sesuai dengan fakta empiris di lapangan. Dengan demikian hipotesis kesatu diterima, artinya model yang tepat untuk mengevaluasi IP-SPM-SD adalah bila di dukung oleh; KLS). PKS, KSK, dan PPOS.

Uji Model Pengukuran



Gambar 3. Pengujian Model Pengukuran

Hasil uji pengukuran menunjukkan bahwa semua variabel terlihat memiliki t-hitung > 2 sehingga dinyatakan signifikan. Hal ini berarti IP-SPM SD berhasil direfleksikan variabel KLS, PKS, KSK, dan PPOS. Maknanya adalah bahwa variabel KLS, PKS, KSK, dan PPOS merupakan variabel yang bisa diangkat sebagai dimensi untuk mengukur IP-SPM-SD. Hasil uji pengukuran dapat dilihat dalam Tabel 1.

Tabel 1. Ringkasan Hasil Pengukuran IP-SPM-SD

No	Variabel	λ Standar	t-hitung	Keterangan*
1	KLS	0.3454	2.5852	Signifikan
2	PKS	0.3103	2.4738	Signifikan
3	KSK	0.2480	2.1298	Signifikan
4	PPOS	0.3929	2.3236	Signifikan

Uji Model Struktural

Berdasarkan nilai t-hitung dalam fungsi struktural bisa dianalisis bahwa:

Pertama, PPOS signifikan terhadap IP-SPM-SD yang ditunjukkan dengan nilai t-hitung sebesar 2.32 > 2. Hasil ini memberi arti bahwa hipotesis kedua dapat diterima. Maknanya adalah bahwa PPOS merupakan

pendukung langsung yang signifikan terhadap IP-SPM-SD. Kedua, untuk menguji hipotesis ketiga, keempat, dan kelima, ada dua fungsi dalam model yang diuji: (1) menjelaskan dukungan KLS, PKS, dan KSK terhadap PPOS, (2) menjelaskan dukungan KLS, PKS, KSK, dan PPOS terhadap IP-SPM SD. Untuk menganalisisnya maka berdasarkan hasil pengujian struktural model dapat dilihat dalam Tabel 2 dan Gambar 4.

Tabel 2. Hasil Pengujian Struktural

Fungsi	Endogen	Eksogen	Koefisien	t-hitung*
1	PPOS	KLS	0.4644	3.4448
		PKS	0.4254	3.1871
		KSK	0.3721	2.8414
2	IP-SPM-SD	KLS	0.3454	2.5852
		PKS	0.3103	2.4732
		KSK	0.2480	2.1298
		PPOS	0.3929	2.3236

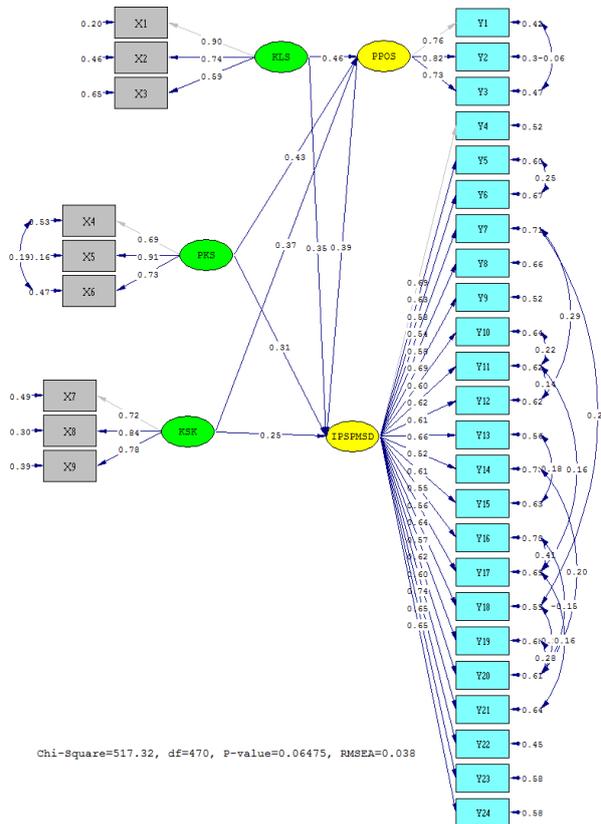
* (Ferdinand, 2000), t-hitung > 2 signifikan

Persamaan fungsi pertama dapat ditulis: $PPOS = 0.4644KLS^{sig} + 0.4254PKS^{sig} + 0.3721KSK^{sig}$. Dalam persamaan ini variabel KLS, PKS, KSK semua signifikan yang berarti bahwa semua variabel tersebut memiliki kontribusi mendukung PPOS. Berdasarkan uraian pada persamaan fungsi pertama tersebut, mengindikasikan bahwa hipotesis ketiga diterima. Maknanya adalah Proses pengelolaan organisasi sekolah didukung langsung secara signifikan oleh kondisi lingkungan sekolah, pelaksana kebijakan sekolah, dan kelompok sasaran kebijakan sekolah.

Persamaan fungsi kedua dapat ditulis, $IP-SPM-SD = 0.3454KLS^{sig} + 0.3103PKS^{sig} + 0.2480KSK^{sig} + 0.3929PPOS^{sig}$. Dalam persamaan semua variabel signifikan yang berarti bahwa variabel KLS, PKS, KSK, dan PPOS memiliki kontribusi dalam menjelaskan IP-SPM-SD. Berdasarkan pada uraian persamaan fungsi kedua ini mengindikasikan bahwa hipotesis keempat diterima. Maknanya adalah bahwa, IP-SPM-SD didukung langsung secara signifikan oleh KLS, PKS, KSK, dan PPOS.

Selain menunjukkan kebermaknaan pengaruh variabel eksogen terhadap variabel endogen, hasil pengujian struktural juga

menunjukkan pengaruh KLS, PKS, dan KSK, terhadap IP-SPM-SD secara tidak langsung melalui PPOS. Dengan demikian hal ini mengindikasikan bahwa hipotesis kelima diterima. Makna hasil pengujian ini adalah bahwa, KLS, PKS, dan KSK merupakan pendukung tidak langsung secara signifikan terhadap IP-SPM-SD. Melalui PPOS.



Gambar 4. Pengujian Model Struktural

Pembahasan

Temuan penelitian ini mendukung teori model implementasi kebijakan yang dikemukakan oleh beberapa ahli: Grindle (1980) yang menyatakan bahwa setiap implementasi kebijakan perlu mempertimbangkan konteks atau kondisi lingkungan dimana tindakan (*action*) dilakukan. Mengacu pada pendapat Grindle berarti bahwa lingkungan adalah merupakan faktor penentu kinerja implementasi kebijakan. Mazmanian & Sabatier (1986) mengemukakan bahwa suatu implementasi kebijakan akan efektif apabila pelaksanaannya mematuhi apa yang telah digariskan oleh peraturan (petunjuk teknis, petunjuk pelaksanaan). Selanjutnya diutarakan

bahwa kelompok sasaran juga harus mematuhi program, tanpa kepatuhan mereka tujuan kebijakan tidak akan tercapai. Mengacu pada pendapat Mazmanian dan Sabatier berarti bahwa kelompok sasaran dan pelaksana kebijakan juga merupakan faktor penentu kinerja implementasi kebijakan.

Van Meter & Van Horn (1973) mengemukakan bahwa kinerja implementasi kebijakan akan ditentukan oleh faktor; standar dan sasaran, sumberdaya, komunikasi antarorganisasi, karakteristik organisasi pelaksana, dan kondisi sosial, ekonomi dan politik. Kelima faktor ini akan membentuk sikap pelaksana terhadap kebijakan yang akan diimplementasikan dan akhirnya mempengaruhi terhadap kinerja implementasi sebuah kebijakan.

Cheema & Rondinelli (1983) mengutarakan adanya 4 faktor yang dipandang dapat mempengaruhi implementasi kebijakan yang disarankan sebagai variabel bebas yakni; kondisi lingkungan, hubungan antarorganisasi, sumber daya, dan karakter agen pelaksana. Membuktikan pendapat Cheema dan Rondinelli, hasil penelitian Khozin (2010, pp. 50–54) bahwa kendala utama pelaksanaan SPM di lapangan adalah pada faktor pendanaan, koordinasi antarunsur dalam organisasi pelaksana dan rendahnya komitmen SDM.

Implikasi diterimanya model teoritis ini diharapkan menjadi *follow up* bagi pengelola program untuk melakukan pembinaan terhadap kinerja sekolah dalam mengimplementasikan kebijakan SPM SD. Implikasi ini di dukung hasil penelitian Rahayu (2015, pp. 62–79) bahwa diperlukan dukungan bagi sekolah-sekolah untuk dapat mengoptimalkan pelaksanaan aspek-aspek pengelolaan sekolah agar lebih sesuai dengan standar-standar pengelolaan pendidikan yang berlaku.

Simpulan

Berdasarkan hasil penelitian dan pembahasan yang disajikan, penelitian ini memberikan simpulan bahwa KLS, PKS, KSK, dan PPOS merupakan faktor-faktor dinamis yang membentuk model yang tepat untuk mengevaluasi IP-SPM-SD.

Daftar Pustaka

- Adawiah, R. (2011). *Standar pelayanan pendidikan di Kabupaten Balangan*. Banjarmasin. Retrieved from <http://eprints.ulm.ac.id/72/>
- Cheema, G. S., & Rondinelli, D. A. (1983). *Policy implementation in developing countris, UNCRD*. Baverly Hill, CA: Sage Publication.
- Dunn, W. N. (1998). *Pengantar analisis kebijakan publik (Terjemahan Samodra Wibawa, dkk)* (2nd ed.). Yogyakarta: UGM Press.
- Edwards III, G. C. (1980). *Implementing public policy*. Washington, D.C: Congressional Quarterly.
- Ferdinand, A. T. (2000). *Structural equation modeling dalam penelitian manajemen*. Semarang: Badan Penerbit Universitas Diponegoro.
- Ghozali, I., & Fuad. (2005). *Structural equation modeling. Teori, konsep dan aplikasi dengan program Lisrel 8.80*. Semarang: Badan Penerbit Universitas Diponegoro.
- Grindle, M. S. (1980). *Politics and policy Implementation in the third world*. New Jersey: Princeton University Press.
- Hair, J. F., Tatham, R. L., Anderson, R. E., & Black, W. (1998). *Multivariate data analysis* (5th ed.). London: Prentice - Hall International.
- Hermawan, K. D. (2011). Evaluasi Program SMP Standar Nasional Berdasarkan Standar Nasional Pendidikan. *Jurnal Pendidikan Dan Kebudayaan*, 17(6), 619. <https://doi.org/10.24832/jpnk.v17i6.54>
- Isaac, S., & Michael, W. B. (1981). *Handbook in research and evaluation for education and behavioral Sciences* (2nd ed.). San Diego, California 92107: EdITS Publishers.
- Jones, C. O. (1994). *An introduction to study of public policy*. Monterey-California: Brooks/Cole Publishing Company.
- Joreskog, K. G., & Sarbon, D. (1996). *Lisrel 8: User's reference guide*. Chicago: Scientific Software International.
- Khozin, M. (2010). Evaluasi implementasi kebijakan standar pelayanan minimal bidang kesehatan di Kabupaten Gunungkidul. *Jurnal Studi Pemerintahan*, 1(1). Retrieved from <http://journal.umy.ac.id/index.php/jspp/article/view/177>
- Koster, W. (2000). Restrukturisasi penyelenggaraan pendidikan: studi kapasitas sekolah dalam rangka desentralisasi pendidikan. *Jurnal Pendidikan Dan Kebudayaan*, 6(26).
- Mazmanian, D., & Sabatier, P. (1986). *Effective olicy implementation*. Lexington: D.C, Heath.
- Menteri Pendidikan dan Kebudayaan Republik Indonesia. (2013). Permendikbud No 23 Tahun 2013 tentang Standar Pelayanan Minimal (SPM) Pendidikan Dasar di Kabupaten/Kota.
- Muhadjir, N. (2007). *Metodologi keilmuan. paradigma kualitatif, kuantitatif, dan mixed* (5th ed.). Yogyakarta: Rake Sarasin.
- Mustopadidjaja, A. R. (2002). *Manajemen proses kebijakan publik. Formulasi, implementasi, dan evaluasi kinerja*. Jakarta: Lembaga Administrasi Negara-RI.
- Narimawati, U., & Sarwono, J. (2007). *Structural equation model (SEM), dalam riset ekonomi menggunakan LISREL*. Yogyakarta: Gava Media.
- Rahayu, M. (2015). Pelaksanaan standar pengelolaan pendidikan di sekolah dasar kecamatan Ngemplak. *Jurnal Penelitian Ilmu Pendidikan*, 8(1), 62–79.
- Rossi, P. H., & Freeman, H. E. (1982). *Evaluation a systematic approach*. London: Sage Publication.
- Sabdaningtyas, L. (2010). *Model evaluasi implementasi kebijakan desentralisasi pendidikan pada satuan sekolah di Propinsi Lampung*. Disertasi Tidak

dipublikasikan.

Satori, D. (1999). *Analisis kebijakan dalam konteks desentralisasi dan otonomi pendidikan*. Jakarta: Biro Perencanaan Sekretariat Jendral Depdikbud.

Solongi, B. (2015, April 11). UN atau Pemenuhan SPM Pendidikan. *Tribunnews*. Retrieved from <http://makassar.tribunnews.com/2015/04/06/un-atau-pemenuhan-spm-pendidikan>

Sudjana, D. (2006). *Evaluasi program pendidikan luar sekolah*. Bandung: PT Remaja Rosdakarya.

Van Meter, D. S., & Van Horn, C. E. (1973). *The policy implementation process a conceptual framework in administration and society*. Beverly Hills: Sage Publication.

Wibawa, S., Prabuningrat, Y., & Pramusinto, A. (1994). *Evaluasi kebijakan publik*. Jakarta: Raja Grafindo Persada.

Worthen, B. R., & Sanders, J. R. (1973). *Educational evaluation: theory and practice*. Belmont, California: Wadsworth Publishing Company, Inc.

NILAI KEUTAMAAN PENGETAHUAN DAN KEBIJAKSANAAN DALAM KONTEKS PENDIDIKAN KARAKTER BANGSA

Wahyu Hidayat^{1*}, Jamil Bin Ahmad², Mohd. Isa Bin Hamzah²

¹Universiti Kebangsaan Malaysia/IAIN Parepare

²Universiti Kebangsaan Malaysia

¹Jl. Amal Bhakti No.8, Bukit Harapan, Soreang, Parepare, Sulawesi Selatan 91131, Indonesia/
43600 Bangi, Selangor, Malaysia

²43600 Bangi, Selangor, Malaysia

* Corresponding Author. Email: wahyuukm82@gmail.com

Abstrak

Pengetahuan dan kebijaksanaan adalah nilai keutamaan tentang fungsi kognitif. Pengetahuan dapat diimplementasikan dalam peningkatan kualitas hidup manusia. Dalam konteks pendidikan karakter bangsa, nilai kreatif, rasa ingin tahu, dan gemar membaca merupakan faktor kekuatan karakter yang membangun nilai keutamaan pengetahuan dan kebijaksanaan. Kajian ini bertujuan untuk memeriksa struktur nilai-nilai karakter bangsa yang menjadi bagian dari nilai keutamaan pengetahuan dan kebijaksanaan. Penelitian ini melibatkan 780 responden siswa SMA. Dapatan kajian menunjukkan bahwa terbentuk tiga faktor kekuatan karakter yaitu kreatif, rasa ingin tahu, dan gemar membaca. Ketiga faktor tersebut merupakan faktor yang dapat membangun nilai keutamaan pengetahuan dan kebijaksanaan.

Kata kunci: *pengetahuan dan kebijaksanaan, pendidikan karakter bangsa*

THE EXCELLENCE VALUES OF KNOWLEDGE AND WISDOM IN NATIONAL CHARACTER EDUCATION CONTEXT

Abstract

Knowledge and wisdom are the excellence value of cognitive function. Knowledge can be implemented in improving the quality of human life. In the context of the nation's character education, the values of creativity, curiosity and reading habits are the factors of character strengths that build the excellence value of knowledge and wisdom. This study is aimed at examining the structure of the nation's character values which become part of knowledge and wisdom priorities. The study involved 780 senior high school students as the respondents. The findings of the study show that there are three factors of character strength, including creativity, curiosity, and reading habits. These three factors are the factors that can build the excellence value of knowledge and wisdom.

Keywords: *knowledge and wisdom, national character education*

Permalink/DOI: <http://dx.doi.org/10.21831/pep.v22i1.18103>

Pendahuluan

Isu pendidikan karakter telah muncul sejak tahun 1990-an. Kemunculan isu pendidikan sebagai jalan keluar dari fenomena-fenomena yang melibatkan pelajar. Kajian (Ma et al. (2008, p. 319) menyebutkan bahwa salah satu dampak buruk dari fenomena tersebut adalah tingkah laku remaja yang disebabkan oleh rendahnya kekuatan karakter mereka. Melalui kajian Josephson Institute of Ethics (2006, pp. 1–3) diketahui bahwa terdapat 82% pelajar yang pernah berbohong kepada orang tua dan 60% pelajar pernah menyontek (*cheating*) dalam ujian. Lebih lanjut, hasil kajian Hart & Carlo (2005, p. 224) menyebutkan bahwa 75% masalah remaja disebabkan oleh rendahnya moral.

Beberapa penelitian menyebutkan kontribusi karakter terhadap keberhasilan pelajar. Park, Peterson, & Seligman (2006, pp. 96–97) mengutarakan bahwa orang yang optimis walaupun memiliki kompetensi rendah, ternyata lebih berhasil dalam bidang pekerjaan dibandingkan dengan orang yang pesimis tapi memiliki kompetensi tinggi. Begitu juga pelajar yang senantiasa optimis dalam bidang akademik memberi pengaruh terhadap pencapaian belajarnya (Bevel & Mitchell, 2012, pp. 780–782; Chang, 2011, pp. 491–515).

Pentingnya moral dan karakter generasi muda menyebabkan beberapa negara maju menjadikan pendidikan karakter sebagai program pendidikan kebangsaan. Beberapa negara menggunakan istilah yang berbeda dalam program pendidikan karakter. Di negara-negara barat seperti Amerika Serikat, Kanada dan Inggris pendidikan karakter lebih dikenali dengan istilah pendidikan nilai, sedangkan di sebagian negara Asia pendidikan karakter lebih populer dengan istilah pendidikan moral. Indonesia, awal mulanya menggunakan istilah pendidikan moral kemudian diubah menjadi pendidikan karakter bangsa.

Program pendidikan karakter di Indonesia bermula pada tahun 2010, kemudian dikukuhkan pada Kurikulum 2013. Kementerian Pendidikan Nasional menetapkan

program pendidikan karakter bangsa yang mengandung 18 nilai karakter bangsa yang perlu diajarkan dan diterapkan dalam pengajaran dan pendidikan di sekolah. Nilai karakter tersebut adalah keagamaan, kejujuran, toleransi, kedisiplinan, kerja keras, kreatif, mandiri, demokratis, memiliki rasa ingin tahu, memiliki semangat kebangsaan, cinta tanah air, menghargai prestasi, bersahabat, cinta perdamaian, gemar membaca, peduli pada alam sekitar, peduli sosial dan bertanggung jawab (Pusat Kurikulum Indonesia, 2009, pp. 9–10). Pendidikan karakter merupakan salah satu metode dalam membina akhlak mulia dan kebiasaan baik bagi pelajar (Schwartz, 2008, pp. 1–23).

Nilai adalah satu set tingkah laku wajar yang mengikuti standar nilai yang dijadikan pedoman oleh individu dan masyarakat sebagai nilai kebaikan (Schwartz, 1994, pp. 19–45). Oleh itu, nilai yang diyakini individu dan masyarakat tersebut menjadi dasar dalam pembentukan karakter seseorang (Lickona, 2012, p. 5). Peterson & Seligman (2004, pp. 13–14) menghubungkan kekuatan karakter dengan nilai kebajikan karena kebajikan (*virtues*) dikembangkan berdasarkan kekuatan karakter. Kekuatan karakter berkontribusi dalam mewujudkan potensi dan cita-cita seseorang dalam kehidupan, bermanfaat bagi orang lain, bangsa dan negaranya. Pengetahuan dan kebijaksanaan adalah nilai keutamaan tentang fungsi kognitif. Peterson & Seligman (2004, p. 29) mengemukakan terdapat lima kekuatan yang meliputi pengetahuan dan kebijaksanaan yaitu kreatif, rasa ingin tahu, mencintai pembelajaran, berfikir kritis dan terbuka serta perspektif.

Dalam konteks teori nilai keutamaan dan kekuatan karakter yang dikemukakan oleh Peterson dan Seligman terdapat persamaan antara nilai karakter bangsa dengan nilai keutamaan dan kekuatan karakter.

Nilai keutamaan merupakan teori yang berpandangan tentang sikap, tingkah laku seseorang. Teori ini tidak memandang perbuatan seseorang itu adil, jujur, atau murah hati. Melainkan teori ini memandangkan apakah seseorang itu berbuat atau bersikap

adil, jujur atau murah hati (Park, 2004; Park et al., 2006). Isu utama teori keutamaan adalah menjelaskan tentang karakter apa saja yang membuat seseorang sebagai orang baik berdasarkan nilai moral. Menurut Peterson & Seligman (2004), keutamaan adalah karakteristik utama daripada karakter

Menurut Peterson & Seligman (2004) keutamaan terbentuk daripada kekuatan karakter. Menurut teori ini, terdapat 24 nilai kekuatan karakter yang membentuk enam nilai keutamaan yaitu kreatif, rasa ingin tahu, berfikir terbuka, mencintai pembelajaran, kebijaksanaan, berani, gigih, integritas, vitalitas, cinta, kebaikan, kecerdasan sosial, kewarganegaraan, keadilan, kepemimpinan, pengampunan, kerendahan hati, kehati-hatian, kontrol diri, apresiasi, rasa syukur, optimisme, humor dan spiritualitas.

Teori keutamaan dan kekuatan karakter menjadi dasar dalam pembuatan instrumen yang mengukur karakter yang disebut *Values in Action Inventory of Strengths* (VIA-IS). Instrumen ini juga digunakan untuk mengukur kekuatan karakter remaja yang disebut *Values in Action Inventory Youth* (VIA-Y). Instrumen ini sudah banyak digunakan di banyak negara. Oleh itu, teori keutamaan dan kekuatan karakter boleh disebut teori yang sudah mapan (*established*).

Nilai keutamaan pengetahuan dan kebijaksanaan memiliki beberapa kekuatan yakni kreatif, rasa ingin tahu, cinta akan pembelajaran, pikiran yang kritis dan terbuka, dan perspektif. Maka nilai-nilai karakter bangsa yang bisa dikategorikan sebagai nilai keutamaan pengetahuan dan kebijaksanaan adalah kreatif, rasa ingin tahu dan gemar membaca.

Dalam teori kekuatan karakter disebutkan kreatif akan memikirkan keterbaruan dan cara yang produktif untuk melakukan sesuatu (Peterson & Seligman, 2004). Pembentukan karakter kreatif dapat dicirikan diantaranya penuh inisiatif, banyak ide, dan inovatif.

Kekuatan karakter rasa ingin tahu merupakan ketertarikan, keterbukaan dalam mencari sesuatu yang baru serta motivasi intrinsik seseorang terhadap pengetahuan dan

pengalaman baru (Peterson & Seligman, 2004). Kekuatan ini dapat diidentifikasi dari minat mencari keterbaruan, mencari informasi, dan terbuka terhadap pengalaman baru.

Gemar membaca merupakan kebiasaan yang menjadi perilaku seseorang untuk membaca secara berulang kali dan bisa dilihat seberapa banyak jumlah bacaan yang dibaca, frekuensi membaca serta rata-rata waktu yang dihabiskan membaca (Chettri & Rout, 2013). Kekuatan karakter gemar membaca dapat diidentifikasi dari kebiasaan meluangkan waktu untuk membaca, mendatangi perpustakaan, mendatangi tempat toko-toko buku bacaan.

Kekuatan karakter di berbagai negara berbeda satu sama lain (Park et al., 2006, pp. 118-129). Hal tersebut disebabkan oleh budaya dan sistem nilai yang dianut oleh suatu bangsa dan negara (Shimai, Otake, Park, Peterson, & Seligman, 2006, pp. 311-322). Kekuatan karakter suatu bangsa mengikuti sistem nilai yang dianut oleh bangsa tersebut, sebagaimana beberapa kajian yang mengadaptasi ke dalam bahasa Jerman (Ruch et al., 2010, pp. 2-12) dan bahasa Hindi (Singh & Choubisa, 2010, pp. 65-76).

Meskipun telah tersedia instrumen VIA-IS (*Value in Action-Inventory Strength*) dan VIA-Y (*Value in Action-Inventory Strength Youth*) yang salah satunya mengukur nilai keutamaan pengetahuan dan kebijaksanaan berserta nilai kekuatannya, kedua instrumen tersebut belum mengakomodir secara keseluruhan nilai-nilai karakter bangsa Indonesia sesuai dengan program pendidikan karakter bangsa. Oleh itu, perlu dikembangkan sebuah alat untuk mengukur karakter yang sesuai dengan bangsa Indonesia, khususnya yang mengukur nilai karakter bangsa yang meliputi kreatif, rasa ingin tahu, dan gemar membaca.

Penelitian ini bertujuan untuk menentukan validitas instrumen nilai keutamaan pengetahuan dan kebijaksanaan yang mengakomodir nilai-nilai karakter bangsa. Pengujian validitas instrumen bertujuan untuk melihat sejauh manakah suatu instrumen dapat mengukur apa yang semestinya diukur

(Nunnally & Bernstein, 1994, p. 83); Pallant, 2016, pp. 90–92). Suatu instrumen dikatakan mempunyai keajegan tinggi apabila mempunyai derajat validitas tinggi dan mengukur apa yang semestinya diukur (Cohen & Swerdlik, 2009, pp. 106–107).

Metode Penelitian

Lokasi penelitian dilaksanakan di dua Kabupaten/Kota yaitu Tanah Toraja yang mewakili etnis Tator dan agama Kristen dan Kota Parepare untuk mewakili etnis Bugis dan agama Islam. Responden dalam penelitian ini adalah siswa SMA di Tanah Toraja dan Parepare sebanyak 780 orang dengan perincian 480 perempuan dan 300 laki-laki. Jumlah sampel ini telah memenuhi persyaratan analisis faktor berdasarkan pendapatnya Hair, Anderson, Tatham, & Black (2010, p. 101) dan Pallant, (2016, p. 178).

Validitas isi dilakukan dengan meminta penilaian pakar psikologi, kewarganega-

raan, agama Islam, agama Kristen dan bimbingan konseling. Skala pengukuran item menggunakan adalah skala Likert yaitu 1 untuk “sangat tidak menggambarkan keadaan saya”, 2 untuk “tidak menggambarkan keadaan saya”, 3 untuk “netral”, 4 untuk “menggambarkan keadaan saya” dan 5 untuk “sangat menggambarkan keadaan saya”. Item-item yang dikembangkan berkaitan dengan nilai kreatif, rasa ingin tahu, dan gemar membaca.

Data yang terkumpul kemudian dianalisis dengan menggunakan analisis faktor eksploratori (EFA) dan analisis faktor konfirmatori (*Confirmatory Factor Analysis/CFA*). Analisis faktor eksploratori bertujuan untuk melihat struktur faktor-faktor tertentu (Hair et al., 2010, p. 102). CFA digunakan untuk menentukan kecocokan model hipotesis dengan metode persamaan pemodelan struktur. Analisis dilakukan dengan bantuan *software* SPSS Versi 23 dan AMOS Versi 20.

Tabel 1. Item-Item Nilai Keutamaan Pengetahuan dan Kebijakan Karakter Bangsa

Dimensi	Ciri	Item	Pernyataan
Kreatif	Inovatif	Krea1	Saya seringkali memikirkan cara-cara lain yang berbeda dan baru dalam melakukan sesuatu
	Banyak idea Inisiatif	Krea2	Saya orang yang memiliki banyak ide
		Krea3	Saya berinisiatif untuk menjawab apa yang ditanyakan guru kepada murid di kelas
Rasa Ingin Tahu	Ketertarikan mencari sesuatu yang baru Mencari informasi tambahan Terbuka terhadap pengalaman baharu	RIT1	Saya sangat senang mencari materi-materi bahan pelajaran yang diminta guru
		RIT2	Sekalipun guru sudah menjelaskan di kelas saya tetap mencari informasi yang lebih detail
		RIT3	Saya membuka diri terhadap perkembangan ilmu baharu
Gemar membaca	Menyediakan waktu untuk membaca Mengunjungi perpustakaan Mendatangi toko buku bacaan	GM1	Saya senantiasa meluangkan waktu untuk membaca
		GM2	Saya rutin mendatangi perpustakaan untuk membaca
		GM3	Saya menyempatkan diri pergi ke toko buku bacaan untuk melihat buku-buku bacaan terbaru

Tabel 2. Analisis Reliabilitas Nilai Alpha Cronbach Untuk Nilai Keutamaan Pengetahuan dan Kebijakan

Aspek Nilai Karakter	Nilai Alpha Cronbach (Sebelum item digugurkan)	Jumlah Item yang digugurkan berdasarkan <i>corrected item total correlation</i> < 0.30	Nilai Alpha Cronbach (sesudah digugurkan)
Kreatif	0,756	-	-
Rasa Ingin Tahu	0,705	2 item	0.721
Gemar Membaca	0,596	2 item	0.700

Hasil Penelitian dan Pembahasan

Sebelum dilakukan analisis faktor eksploratori (EFA) dan analisis faktor konfirmatori (CFA) diperlukan uji persyaratan, salah satunya adalah uji normalitas. Anderson & Gerbing (1988, pp. 411–423) yang mempersyaratkan data berdistribusi normal pada peringkat univariat adalah diantara + 2 hingga -2. Secara umum nilai z skewness dan kurtosis berdistribusi secara normal. Hair et al. (2010, p. 97) menyebutkan data berdistribusi normal jika semua item dari sebuah instrumen berada pada sebuah garis lurus sudut 45 derajat. Analisis faktor eksploratori pada kajian ini merujuk pada persyaratan yang dikemukakan pakar, sebagaimana Tabel 3. Persyaratan analisis faktor eksploratori mengikut saran (Hair et al. (2010).

Tabel 3. Prasyarat Analisis Faktor Eksploratori

Bil	Kriteria	Kriteria
1	Ujian <i>Kaiser-Meyer-Olkin</i> (KMO)	0,5-0,7
2	Ujian <i>Bartlett's Test of Sphericity</i>	< 0,05
3	<i>Anti-image correlation matrix of items</i>	≥ 0,05
4	Nilai <i>variables communalities</i>	≥ 0,05
5	Nilai <i>Loading Factor</i>	≥ 0,30
6	Nilai <i>Eigen Factor</i>	> 1,00
7	Persentase <i>Variance</i>	≥ 60%

Analisis faktor merupakan teknik untuk menentukan validitas konstruk instrumen. Dalam kajian ini peneliti menjalankan analisis faktor untuk setiap aspek nilai dalam karakter bangsa. Untuk tujuan *exploratory*, penelitian ini menggunakan teknik komponen prinsipal (*principal component*) orthogonal dengan putaran varimax. Putaran varimax dipilih berdasarkan persyaratan item-item dari konstruk yang berbeda tidak mempunyai korelasi atau hubungan yang baik dan juga metode ini mudah ditafsirkan, sebaliknya putaran oblimax tidak dipilih karena mempersyaratkan item-item atau faktor-faktornya mempunyai korelasi (Pallant, 2016, pp. 182–183). Menurut Pallant (2016, pp. 187–190), putaran varimax bertujuan mengurangi bilangan faktor yang kompleks melalui peningkatan varians yang dimuatkan pada setiap faktor.

Pengkaji juga mengambil nilai pemberat (*weight value*) 0.6 dalam penilaian bilangan sampel mengikut kaedah Keiser-Meyer-Olkin's (KMO). Menurut Pallant (2016, p. 182), nilai > 0.60 adalah mencukupi. Hasil penelitian menunjukkan nilai KMO yang diperoleh adalah 0,801. Hal tersebut berarti telah memenuhi kriteria minimum yang disarankan pakar. Selanjutnya, nilai *Barlett's Test of Sphericity* adalah signifikan secara statistik ($p < 0,05$). Hal ini menunjukkan bahwa secara keseluruhan korelasi menggunakan *correlation matrix* adalah signifikan.

Berikut ini nilai kecukupan sampel mengikut kaidah KMO untuk nilai keutamaan pengetahuan dan kebijaksanaan.

Tabel 4. Kecukupan Sampel Analisis Faktor Eksploratori (*Exploration Factor Analysis*)

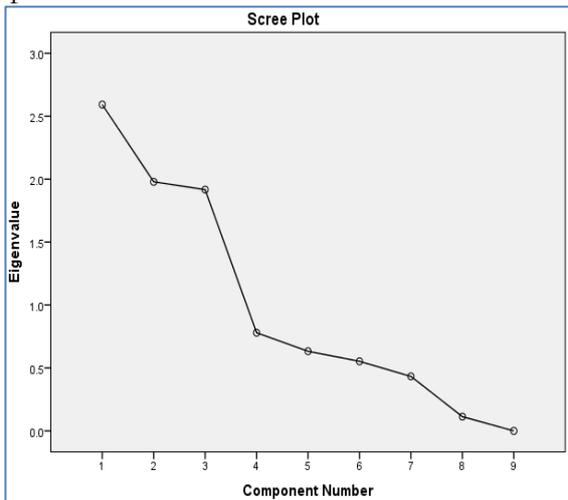
Kaiser-Meyer-Olkin Measure of Sampling Adequacy.	0,799
Bartlett's Test of Sphericity	Approx. Chi-Square 1233,008
	Df 36
	Sig. 0,000

Tabel 5. Matriks Komponen dengan Putaran (Varimax Variabel) Nilai Keutamaan Pengetahuan dan Kebijaksanaan

Ciri-Ciri	Faktor 1	Faktor 2	Faktor 3
Inovatif			0,791
Banyak idea			0,850
Inisiatif			0,792
Ketertarikan	0,772		
Mencari informasi	0,619		
Terbuka dengan pengalaman	0,672		
Menyediakan waktu untuk membaca		0,728	
Mengunjungi perpustakaan		0,640	
Mengunjungi toko buku bacaan		0,742	
Eigen	2,304	2,205	1,979
% of Variance	25,601	24,497	21,991
Cumulative %	25,601	50,099	72,090
Nilai Karakter Bangsa	Gemar Membaca	Rasa Ingin Tahu	Kreatif

Tabel 5 menunjukkan pembentukan faktor nilai keutamaan pengetahuan dan kebijaksanaan dalam konteks pendidikan karakter bangsa. Gemar membaca adalah faktor pertama, rasa ingin tahu sebagai faktor kedua dan kreatif sebagai faktor ketiga.

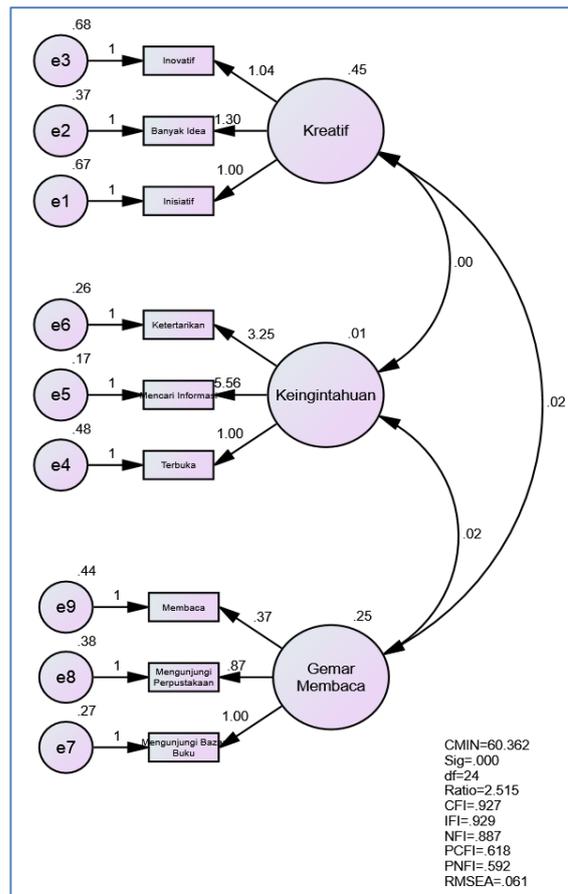
Diagram *scree plot* pada Gambar 1 menunjukkan bagaimana kecenderungan penurunan nilai *eigen* (*eigenvalues*) yang dipakai untuk menentukan banyaknya faktor yang dipakai.



Gambar 1. *Scree plot* pada ujian pengesahan

Mengacu pada pendapat Tabachnick & Fidell (2013, p. 709) suatu model dikatakan memenuhi kriteria model yang sesuai jika memenuhi sekurang-kurangnya satu indeks model fit. Indeks model fit yang dimaksud adalah tes Chi-square (χ^2), *Root Mean Square of Error Approximation* (RMSEA), *GFI (Goodness of Fit Index)*, *AGFI (Adjusted Goodness of Fit Index)*, *CFI* dan *TLI*. Untuk mencapai kesesuaian model, nilai relatif Chi-square semestinya kurang dari 5,0 sedangkan *CFI* dan *TLI* lebih dari 0,9. Untuk nilai *RMSEA* hendaknya kurang dari 0,08 (Schumacker & Lomax, 2004, p. 261). Gambar 2 menunjukkan model CFA pada penelitian ini.

Pada tabel 6 ditunjukkan kriteria model fit CFA nilai keutamaan pengetahuan dan kebijaksanaan pada karakter bangsa.



Gambar 2. Model CFA Nilai Keutamaan Pengetahuan dan Kebijakan Karakter Bangsa

Tabel 6. Analisis Faktor Konfirmatori (*Confirmatory Factor Analysis*) (IPKB)

Goodness of fit Indexs	Cut off Value	Hasil Model	Keterangan
χ^2 - Chi square	-	60,362	
Probability	$\geq 0,05$	0,000	Kurang Baik
RMSEA	$\leq 0,08$	0,061	Baik
CFI	$\geq 0,90$	0,927	Baik
AGFI	$\geq 0,90$	0,943	Baik
TLI	$\geq 0,90$	0,891	Kurang Baik
GFI	$\geq 0,90$	0,970	Baik

Tabel 6 menunjukkan hasil analisis faktor konfirmatori nilai keutamaan pengetahuan dan kebijaksanaan. Nilai indeks model fit adalah bersamaan dengan $p=0,000$, nilai relatif tes Chi-square secara bersamaan adalah 60.362, nilai 0,927 untuk *CFI*, nilai 0,943 untuk *AGFI*, nilai 0,891 untuk *TLI*,

nilai 0,970 untuk GFI dan 0,061 untuk RMSEA.

Tabel 7. Validitas Konstruksi Nilai Keutamaan Pengetahuan dan Kebijakan Karakter Bangsa

Dimensi	Item	AVE	CR
Kreatif	Krea1	0,701	0,859
	Krea2		
	Krea3		
Keingintahuan	RIT1	0,501	0,782
	RIT2		
	RIT3		
Gemar Membaca	GM1	0,515	0,790
	GM2		
	GM3		

Keterangan:

AVE = Average Variance Extracted

CR = Composite Reliability

Menurut Hair et al. (2010) Hair et al. (2010), reliabilitas perlu memenuhi tiga kriteria yaitu reliabilitas internal yang ditentukan oleh nilai Alpha Cronbach lebih dari 0,70, nilai reliabilitas konstruk atau *composite reliability* (CR) lebih dari 0,60 dan nilai *average variance extracted* (AVE) melebihi 0,40 sebagaimana ditunjukkan pada Tabel 6.

Tabel 7 menunjukkan semua konstruk telah memenuhi kriteria reliabilitas yaitu nilai AVE lebih 0,40 dan nilai CR lebih dari 0,70. Ini menunjukkan bahwa model pengukuran tahap validitas dan reliabilitas mempunyai kriteria yang baik.

Tujuan penelitian ini adalah untuk memeriksa struktur nilai-nilai karakter bangsa yang menjadi bagian dari nilai keutamaan pengetahuan dan kebijakan. Penelitian ini telah berhasil membuktikan bahwa terdapat tiga faktor nilai karakter bangsa yakni, kreatif, rasa ingin tahu dan gemar membaca yang dapat menjadi bagian dari pembentuk nilai keutamaan pengetahuan dan kebijakan.

Berdasarkan teori keutamaan dan kekuatan karakter Peterson dan Seligman, nilai keutamaan pengetahuan dan kebijakan terbentuk dari kekuatan karakter kreatif, rasa ingin tahu, mencintai pembelajaran, pikir-

an terbuka dan prespektif. Persamaan kekuatan karakter untuk nilai keutamaan pengetahuan dan kebijakan dalam konteks pendidikan karakter bangsa adalah kreatif, rasa ingin tahu dan gemar membaca.

Penelitian ini serupa dengan beberapa kajian yang menunjukkan bahwa kreatif, rasa ingin tahu, mencintai pembelajaran bagian dari nilai keutamaan pengetahuan dan kebijakan (Ruch & Proyer, 2015, pp. 2–12). Kajian Macdonald, Bore, & Munro (2008, pp. 787–799) memasukkan kekuatan karakter kreatif, rasa ingin tahu, mencintai pembelajaran ke dalam faktor intelektual dengan nilai *loading factor* masing-masing -0,78, -0,61 dan -0,60. Kajian Singh & Choubisa (2010, pp. 325–332), Littman-Ovadia & Lavy (2012, pp. 41–50), dan Ruch et al. (2010, pp. 138–149) memasukkan ke dalam faktor *Intellectual Strengths*. Pembentukan kekuatan karakter kreatif, rasa ingin tahu dan gemar membaca dalam nilai keutamaan kebijakan dan pengetahuan sesuai dengan Peterson dan Seligman.

Simpulan

Berdasarkan hasil penelitian dan pembahasan dapat disimpulkan bahwa dalam nilai keutamaan kebijakan dan pengetahuan konteks pendidikan karakter terbentuk tiga nilai kekuatan karakter yaitu nilai karakter gemar membaca, rasa ingin tahu dan kreatif. Item-item dalam kajian ini telah sesuai dengan ciri-ciri yang dikemukakan dalam pengembangan instrumen ini.

Daftar Pustaka

- Anderson, J. C., & Gerbing, D. W. (1988). Structural equation modeling in practice: A review and recommended two-step approach. *Psychological Bulletin*, 103(3), 411–423.
- Bevel, R. K., & Mitchell, R. M. (2012). The effects of academic optimism on elementary reading achievement. *Journal of Educational Administration*, 50(6), 773–787.
<https://doi.org/10.1108/09578231211>

264685

- Chang, I.-H. (2011). A study of the relationships between distributed leadership, teacher academic optimism and student achievement in Taiwanese elementary schools. *School Leadership & Management*, 31(5), 491–515. <https://doi.org/10.1080/13632434.2011.614945>
- Chettri, K., & Rout, S. K. (2013). Reading habits - an overview. *IOSR Journal Of Humanities And Social Science (IOSR-JHSS)*, 14(6). Retrieved from <http://www.iosrjournals.org/iosr-jhss/papers/Vol14-issue6/C01461317.pdf>
- Cohen, R. J., & Swerdlik, M. (2009). *Psychological testing and assessment: an introduction to tests and measurement* (7th ed.). USA: McGraw-Hill.
- Hair, J. F., Anderson, R. E., Tatham, R. L., & Black, W. C. (2010). *Multivariate data analysis* (4th ed.). New Jersey: Prentice Hall.
- Hart, D., & Carlo, G. (2005). Moral development in adolescence. *Journal of Research on Adolescence*, 15(3), 223–233. <https://doi.org/10.1111/j.1532-7795.2005.00094.x>
- Josephson Institute of Ethics. (2006). The ethics of American youth. Retrieved from <http://www.josephsoninstitute.org>
- Lickona, T. (2012). *Character Matters: Persoalan Karakter, (terj. Juma Wadu Wamaungu & Jean Antunes Rudolf Zien dan Editor Uyu Wahyuddin dan Suryani)*. Jakarta: Bumi Aksara.
- Littman-Ovadia, H., & Lavy, S. (2012). Character strengths in Israel. *European Journal of Psychological Assessment*, 28(1), 41–50. <https://doi.org/10.1027/1015-5759/a000089>
- Ma, M., Kibler, J. L., Dollar, K. M., Sly, K., Samuels, D., Benford, M. W., ... Wiley, F. (2008). The relationship of character strengths to sexual behaviors and related risks among African American adolescents. *International Journal of Behavioral Medicine*, 15(4), 319–327. <https://doi.org/10.1080/10705500802365573>
- Macdonald, C., Bore, M., & Munro, D. (2008). Values in action scale and the Big 5: An empirical indication of structure. *Journal of Research in Personality*, 42(4), 787–799. <https://doi.org/10.1016/j.jrp.2007.10.003>
- Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York: McGraw Hill.
- Pallant, J. (2016). *SPSS survival manual: a step by step guide to data analysis using IBM SPSS*. Sydney: Allen & Unwin.
- Park, N. (2004). Character strengths and positive youth development. *The ANNALS of the American Academy of Political and Social Science*, 591(1), 40–54. <https://doi.org/10.1177/0002716203260079>
- Park, N., Peterson, C., & Seligman, M. E. P. (2006). Character strengths in fifty-four nations and the fifty US States. *The Journal of Positive Psychology*, 1(3), 118–129.
- Peterson, C., & Seligman, M. E. P. (2004). *Character strengths and virtues: a handbook and classification*. Oxford: Oxford University Press.
- Pusat Kurikulum Indonesia. (2009). *Pengembangan dan pendidikan budaya karakter bangsa*. Jakarta: Kementerian Pendidikan dan Kebudayaan republik Indonesia.
- Ruch, W., & Proyer, R. T. (2015). Mapping strengths into virtues: the relation of the 24 VIA-strengths to six ubiquitous virtues. *Frontiers in Psychology*, 6. <https://doi.org/10.3389/fpsyg.2015.00460>

- Ruch, W., Proyer, R. T., Harzer, C., Park, N., Peterson, C., & Seligman, M. E. P. (2010). Values in action inventory of strengths (VIA-IS). *Journal of Individual Differences*, 31(3), 138–149. <https://doi.org/10.1027/1614-0001/a000022>
- Schumacker, R., & Lomax, R. (2004). *A beginner's guide to structural equation modeling* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum.
- Schwartz, M. J. (2008). Introduction to character education and effective principles. In M. J. Schwartz (Ed.), *Effective character education: A guidebook for future educators* (pp. 1–23). New York: McGraw-Hill.
- Schwartz, S. H. (1994). Are there universal aspects in the structure and contents of human values? *Journal of Social Issues*, 50(4), 19–45. <https://doi.org/10.1111/j.1540-4560.1994.tb01196.x>
- Shimai, S., Otake, K., Park, N., Peterson, C., & Seligman, M. E. P. (2006). Convergence of character strengths in American and Japanese young adults. *Journal of Happiness Studies*, 7(3), 311–322. <https://doi.org/10.1007/s10902-005-3647-7>
- Singh, K., & Choubisa, R. (2010). Empirical validation of values in action-inventory of strengths (VIA-IS) in Indian context. *Psychological Studies*, 55(2), 151–158. <https://doi.org/10.1007/s12646-010-0015-4>
- Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (6th ed.). Boston: Pearson & Bacon.

PENGEMBANGAN INSTRUMEN ASESMEN DIRI GURU TERHADAP AKTIVITAS PENGEMBANGAN KEPROFESIAN BERKELANJUTAN

Wirman Kasmayadi^{1}, Kumaidi², Sumarno³*

¹LPMP Prov. Tenggara Barat, ²Univ. Muhammadiyah Surakarta, ³Univ. Negeri Yogyakarta

¹Jl. Panji Tilar Negara No.8, Taman Sari, Ampenan, Mataram, NTB 83114, Indonesia

²Jl. A. Yani, Mendungan, Pabelan, Kartasura, Kab. Sukoharjo, Jawa Tengah 57162, Indonesia

³Jl. Colombo No. 1, Depok, Sleman 55281, Yogyakarta, Indonesia

* Corresponding Author. Email: wirkasmayadi@yahoo.com

Abstrak

Penelitian ini bertujuan untuk mengembangkan instrumen asesmen diri guru pada aktivitas pengembangan keprofesian berkelanjutan yang handal. Metode penelitian yang digunakan adalah penelitian pengembangan yang mengacu pada lima tahapan dalam model Plomp. Subyek penelitian adalah guru SMA di Provinsi Nusa Tenggara Barat yang dipilih dengan teknik cluster random sampling. Jumlah subyek penelitian pada uji coba awal sebanyak 350 orang dan uji coba lapangan utama sebanyak 950 orang. Pengumpulan data menggunakan kuesioner skala Likers (1-5). Pengujian validitas isi menggunakan expert judgment, sedangkan validitas dan reliabilitas empiris menggunakan analisis faktor konfirmatori (CFA) dengan bantuan program Lisrel 8.80. Hasil penelitian menunjukkan bahwa: (1) konstruk aktivitas pengembangan keprofesian berkelanjutan guru memiliki tujuh dimensi, yaitu: membaca, pengembangan pembelajaran dan penilaian, pendidikan dan pelatihan, publikasi ilmiah, publikasi karya inovatif, reflektif dan kolaboratif; (2) instrumen yang telah dikembangkan secara keseluruhan memiliki kecocokan sebagai model asesmen yang baik dengan RMSEA 0.005 dan CFI 1.00 (valid) dan composite reliability lebih dari 0.70 (reliable).

Kata kunci: *asesmen diri guru, pengembangan keprofesian berkelanjutan*

DEVELOPING TEACHERS' SELF-ASSESSMENT INSTRUMENT FOR SUSTAINABLE PROFESSIONAL DEVELOPMENT ACTIVITIES

Abstract

This research is aimed at developing a reliable teachers' self-assessment instrument for sustainable professional development activity. This study was a development research using five-stage model of Research and Development by Plomp. The subjects are high school teachers in the West Nusa Tenggara province, that were selected by cluster random sampling technique. The research subjects in preliminary field test were 350 people and in the main field test were 950 people. The data were collected using Likers scale (1-5) questionnaire. The content validity was tested using expert judgment, whereas the empirical validity and reliability were tested using confirmatory factor analysis (CFA) with Lisrel 8.80 program. The results of the research show that: (1) the construct of teachers' sustainable professional development activities has seven dimensions, namely: reading, learning and assessment development, education and training, scientific publications, the publication of innovative works, reflective and collaborative; (2) the instruments have an overall fit as a good assessment model with the value of RMSEA 0.005, CFI 1.00 (Valid), and composite reliability of more than 0.70 (reliable).

Keywords: *teachers' self-assessment, sustainable professional development*

Permalink/DOI: <http://dx.doi.org/10.21831/pep.v22i1.8290>

Pendahuluan

Penyelenggaraan pendidikan yang berkualitas membutuhkan guru profesional. Predikat profesional dapat terwujud apabila guru dapat mengembangkan keprofesiannya secara berkelanjutan. Tuntutan terhadap pengembangan keprofesian berkelanjutan semakin tinggi seiring dengan hadirnya dinamika perubahan saat ini, dimana guru dihadapkan pada perubahan yang cepat, permintaan standar yang tinggi, dan tuntutan peningkatan mutu pendidikan.

Pengembangan keprofesian menjadi lebih diperlukan oleh guru sebagai fasilitator kegiatan pembelajaran di sekolah. Guru dan fasilitator memiliki tanggungjawab untuk membelajarkan peserta didik agar mampu menghadapi dinamika perubahan yang sedang dan akan terjadi, baik perubahan teknologi, ilmu pengetahuan, maupun struktur ketenagakerjaan. Pada kondisi demikian, guru harus memiliki inisiatif sendiri dan termotivasi secara intrinsik untuk selalu *update*, meningkatkan dan memperluas kompetensinya melalui pengembangan keprofesian secara berkelanjutan.

Perlunya pengembangan keprofesian berkelanjutan bagi guru didukung oleh beberapa hasil studi. Berdasarkan hasil uji kompetensi guru (UKG), bahwa kompetensi profesional dan pedagogis guru di Indonesia masih rendah, dimana nilai rata-rata nasional UKG 2012 sebesar 44,5 (Zubaidah, 2012) serta UKG 2015 dengan nilai 53,02 (Wurinanda, 2015). Kondisi guru tersebut berimplikasi pada rendahnya pencapaian prestasi peserta didik Indonesia pada tes standar internasional tahun 2011 (TIMSS, PISA, PIRLS) yang menunjukkan bahwa peringkat peserta didik Indonesia berada pada posisi kelompok bawah (Mullis, Martin, Foy, & Arora, 2012, p. 42).

Istilah yang berkaitan dengan pengembangan keprofesian berkelanjutan diantaranya adalah *continuing professional development*. Menurut Kelchtermans (2004, p. 220) bahwa pengembangan keprofesian berkelanjutan merupakan proses pembelajaran dari interaksi dengan lingkungan sosial, sekolah dan budaya, sehingga menimbulkan

perubahan dalam praktik (tindakan) profesional dan pemikiran guru tentang praktiknya. Dengan demikian, pengembangan keprofesian berkelanjutan menunjukkan proses belajar guru dan hasil belajar ini akan terlihat tidak hanya dalam perubahan praktik profesional seseorang (misalnya, performa yang lebih efektif), tetapi juga dalam berpikir tentang bagaimana dan mengapa praktik itu.

Day & Sachs (2004, p. 219) mengungkapkan bahwa pengembangan keprofesian berkelanjutan sebagai upaya pengembangan yang terdiri dari seluruh pengalaman belajar yang dialami dan aktivitas-aktivitas sadar yang direncanakan dan dimaksudkan secara langsung maupun tidak langsung untuk memberi manfaat pada individu, kelompok, sekolah, serta berkontribusi pada kualitas pendidikan di ruang kelas. Dengan demikian, pengembangan keprofesian mencakup segala aktivitas yang memenuhi baik kebutuhan individual guru maupun kebutuhan institusional sekolah. Guru, sekolah, dan peserta didik mendapatkan manfaat dari proses pengembangan keprofesian tersebut.

Definisi ini menyoroti fungsi dan peran pengembangan keprofesian berkelanjutan secara terpisah, namun saling terkait antara aspek fungsi, hubungan antara individu dan kolegialitas dalam kegiatan pengembangan. *Pertama*, pengembangan keprofesian berkelanjutan guru secara personal dapat saja memiliki fungsi yang berbeda, namun tetap berorientasi pada pemeliharaan, peningkatan, atau perluasan kompetensi. Guru harus belajar secara berkelanjutan agar dapat memperbarui pengetahuan dan keterampilannya dengan belajar strategi, teknik, dan metode baru untuk memenuhi tantangan kelas baru, memperbaiki diri dan mengubah sekolah menjadi masyarakat belajar, serta dapat merespon perubahan pendidikan.

Kedua, fungsi pengembangan keprofesian berkelanjutan menekankan adanya hubungan simbiosis yang saling menguntungkan antara kebutuhan individu dan organisasi. Dalam komunitas belajar profesional, guru terfokus pada belajar peserta

didik, berbagi nilai dan visi, tanggung jawab kolektif, inkuiri, profesional reflektif, kolaborasi, serta belajar kelompok dan individu. Berpartisipasi dalam kegiatan profesional yang variatif berarti guru telah menstimulus dirinya sebagai pembelajar mandiri dan pengembangan sekolah, sehingga guru memberikan kontribusi yang signifikan untuk meningkatkan praktik pendidikan.

Gaible & Burns (2005, pp. 15–16) mengklasifikasikan model pengembangan keprofesian guru menjadi tiga kelompok besar, yaitu standar/formal; mandiri/informal; dan berbasis tempat. *Pertama*, pengembangan keprofesian guru standar/formal yang didefinisikan oleh de Vries, Jansen, & van de Grift (2013, p. 870) sebagai kesempatan belajar formal dengan lingkungan belajar terstruktur dengan kurikulum tertentu, penyelenggaraan secara formal dan difasilitasi oleh lembaga atau organisasi pemerintah ataupun lembaga swadaya masyarakat sebagai penyelenggara kegiatan belajar.

Dalam implementasinya, aktivitas pengembangan keprofesian guru standar diselenggarakan dalam bentuk program kualifikasi guru melalui pendidikan gelar atau non gelar. Selain itu, aktivitas belajar formal lainnya, seperti kursus, *workshop*, pelatihan, membaca buku siswa, pedoman guru serta jurnal profesional dalam upaya menyegarkan dan memperbaharui pengetahuan dan keterampilan dengan konten terbaru, terkait dengan materi pelajaran, metode/strategi pembelajaran, pengembangan kurikulum, dan kegiatan sejenis lainnya.

Kedua, pengembangan keprofesian guru mandiri atau informal. Menurut Desimone (2009, p. 287), pengembangan keprofesian informal tidak mengikuti kurikulum yang ditentukan dan tidak terbatas pada lingkungan tertentu. Aktivitas belajar meliputi kegiatan individu seperti membaca buku dan observasi kelas serta kegiatan kolaboratif seperti diskusi dengan kolega dan orang tua, mentoring, jaringan guru dan kelompok penelitian. Partisipasi dalam kegiatan belajar informal, umumnya tidak wajib, tetapi atas inisiatif guru sendiri. Oleh karena itu, guru tidak hanya sebagai pe-

nerima pengetahuan, tetapi sebagai pembelajar mandiri dengan mengatur proses belajar dan menentukan tujuan belajar serta strategi secara mandiri.

Selain itu, kesempatan belajar informal sering tertanam dalam kelas atau konteks sekolah, yang memungkinkan para guru untuk merefleksikan praktik dan belajar dari koleganya (Borko, 2004, p. 256). Dengan demikian, dapat dijelaskan bahwa pengembangan keprofesian informal atau mandiri sangat cocok bagi guru dengan alasan, yaitu: (1) karakteristik guru sebagai pembelajar orang dewasa (*adult learner*), (2) kesempatan belajar yang lebih besar, ditinjau dari perspektif rencana, inisiatif, proses dan strategi/metode, serta pelaksanaannya yang dikendalikan sendiri oleh pembelajar, dan (3) pilihan kegiatan sangat variatif sesuai dengan kebutuhan guru dan tidak terbatas pada pilihan aktivitas yang secara langsung terkait dengan kebutuhan kelas maupun kinerja profesinya.

Ketiga, pengembangan keprofesian guru berbasis-tempat. Day & Sachs (2004, p. 13) membedakan pengembangan keprofesian guru meliputi, (1) belajar di sekolah seperti melalui: *coaching* teman sejawat, *mentoring*, *critical friendship*, *active research*, *team planning and assessment*, *appraisal*; (2) belajar di luar sekolah seperti melalui: *school-led renewal or reform networks*, *school-university partnership*, *subject or phone-specific network*, *professional development centres*.

Sehubungan dengan guru sebagai pembelajar mandiri (*autonomous learner*), Benson (2007, p. 345) berpendapat bahwa pembelajar yang mandiri merupakan pembelajar yang mampu mengambil tanggung jawab untuk menentukan tujuan, konten, dan metode pembelajaran mereka sendiri, serta mampu memonitor kemajuannya dan mengevaluasi luaran yang dihasilkan. Dengan demikian, para guru seharusnya berperilaku sebagai pembelajar mandiri dalam pengembangan keprofesian.

Terkait dengan konteks bekerja (sekolah) sebagai *setting* pelaksanaan aktivitas pengembangan keprofesian guru, ditemukan beberapa hasil penelitian. *Pertama*, Kwakman

(2003, pp. 159–160) melakukan studi literatur terhadap aktivitas belajar keprofesian guru, hasilnya membedakan empat kategori aktivitas belajar guru, yaitu: (1) membaca, (2) bereksperimen, (3) refleksi, dan (4) kolaborasi.

Kedua, Hoekstra, Brekelmans, Beijaard, & Korthagen (2009, pp. 664–665) mengungkapkan empat kategori utama dari kegiatan belajar guru di tempat kerja, yaitu: (1) belajar dengan bereksperimen, (2) belajar dengan mempertimbangkan praktik mengajar sendiri, (3) belajar dengan mendapatkan gagasan dari kolega, dan (4) belajar sambil bekerja.

Ketiga, Bakkenes, Vermunt, & Wubbels (2010, pp. 539–540) telah melakukan analisis konten terhadap pengalaman belajar guru dan hasilnya diklasifikasikan menjadi enam kategori, yaitu (1) eksperimen (mencoba sesuatu yang baru dalam praktik pembelajaran); (2) mempertimbangkan praktik sendiri (refleksi terhadap praktik pembelajaran dan hasil belajar peserta didik); (3) mendapatkan ide dari orang lain; (4) memperhatikan kesenjangan antara harapan dengan realitas di kelas; (5) berusaha untuk mengubah pengajaran agar tidak kembali ke rutinitas lama; dan (6) terlibat dalam aktivitas pengembangan agar terus belajar tentang metode/strategi baru.

Keempat, Tynjälä (2008, p. 135) telah melakukan studi literatur dan melakukan kajian bagaimana individu belajar di tempat kerja, yaitu: (1) melakukan pekerjaan sendiri; (2) berinteraksi dengan kolega; (3) bekerja dengan pelanggan; (4) mengerjakan tugas yang menantang dan tugas baru; (5) melakukan refleksi dan evaluasi pengalaman kerja; (6) pendidikan formal; dan (7) melakukan kerja ekstra.

Kelima, pelaksanaan pengembangan keprofesian berkelanjutan guru di Indonesia mengacu pada Pedoman Pengelolaan Pengembangan Keprofesian Berkelanjutan (PKB) Guru (Menteri Pendidikan dan Kebudayaan Republik Indonesia, 2010, p. 18) dengan aktivitas guru meliputi: (1) kegiatan mandiri di sekolah, (2) kegiatan di sekolah dengan berkolaborasi bersama guru

lain, (3) kegiatan melalui jaringan sekolah, dan (4) kegiatan yang difasilitasi oleh kelompok kepakaran lain, seperti dosen, widyaiswara, dan guru inti.

Berdasarkan hasil kajian literatur dan mengacu pada aktivitas pengembangan keprofesian berkelanjutan guru di Indonesia, maka dimensi aktivitas pengembangan keprofesian berkelanjutan guru pada penelitian ini, adalah: (1) membaca, (2) pengembangan pembelajaran dan penilaian, (3) pendidikan dan pelatihan, (4) publikasi ilmiah, (5) publikasi karya inovatif, (6) refleksi, dan (7) kolaborasi.

Hasil kajian menunjukkan bahwa asesmen diri merupakan cara ampuh untuk mengembangkan otonomi dalam belajar. Namun pada kenyataannya, kajian sistematis mengenai hal ini belum banyak dilakukan oleh peneliti pendidikan. Demikian pula pengamatan di lapangan menunjukkan bahwa para pendidik belum secara sadar mengimplementasikan asesmen diri dalam pengembangan keprofesian berkelanjutan.

Asesmen diri merupakan pelibatan pembelajar dalam menentukan standar dan/atau kriteria keberhasilan untuk menilai hasil belajarnya sendiri, sehingga dapat menentukan sejauhmana keberhasilan telah tercapai berdasarkan standar atau kriteria (Brew, 1999, p. 568). Dengan demikian, terdapat dua unsur pada setiap proses asesmen, yaitu penentuan standar terkait dengan kriteria tertentu, dan penilaian terhadap karya berdasarkan standar dan kriteria. Melalui asesmen diri, pembelajar dapat melihat, bukan hanya kekurangannya, tetapi juga kelebihan yang dimilikinya, untuk selanjutnya kekurangan ini akan menjadi tujuan perbaikan, sedangkan kelebihannya akan menjadi motivasi untuk selalu berkembang.

Sehubungan konsep asesmen diri, Rolheiser & Ross (2005, p. 254) telah mengajukan suatu model teoretik untuk menunjukkan kontribusi asesmen diri pada pencapaian tujuan belajar. Model ini menekankan pada proses evaluasi diri. Saat mengevaluasi diri sendiri, pembelajar terdorong untuk menetapkan tujuan yang lebih tinggi (*goals*) dan berusaha lebih keras (*effort*).

Kombinasi tujuan *goals* dan *effort* ini menentukan prestasi (*achievement*). Selanjutnya prestasi ini berakibat pada penilaian diri (*self-judgment*) melalui kontemplasi seperti pertanyaan, “apakah tujuanku telah tercapai”? akibatnya timbul reaksi (*self-reaction*), seperti “apa yang dirasakan dari prestasinya?”. Dengan demikian, maka *goals*, *effort*, *achievement*, *self-judgment*, dan *self-reaction* dapat terpadu untuk membentuk kepercayaan diri (*self-confidence*) yang positif. Selanjutnya, kedua penulis juga menekankan bahwa asesmen diri adalah kombinasi dari komponen *self-judgment* dan *self-reaction* dalam model di atas.

Selain itu, kedua penulis di atas meyakini bahwa asesmen diri dapat meningkatkan hasil belajar, karena beberapa alasan, yaitu: (a) dapat memusatkan perhatian pebelajar pada tujuan pembelajaran, (b) memberikan informasi pada guru mengenai hal-hal yang masih kurang atau belum tercapai dalam pembelajaran, (c) dapat lebih meningkatkan perhatian pebelajar pada asesmen, dan (d) meningkatkan motivasi pebelajar.

Melalui asesmen diri, pebelajar akan dapat menentukan tujuan belajarnya dan berusaha keras untuk mencapai tujuan itu. Oleh karena itu, asesmen diri secara langsung dapat meningkatkan proses dan hasil belajar, sehingga mereka mampu mengarahkan dirinya untuk belajar secara mandiri.

Terkait dengan kegiatan pengukuran aktivitas pengembangan keprofesian berkelanjutan, diperlukan suatu instrumen untuk melaksanakan asesmen diri guru terhadap aktivitas pengembangan keprofesian berkelanjutan yang telah teruji, baik validitas maupun reliabilitasnya. Dalam kegiatan penelitian khususnya dalam bidang pendidikan, terdapat dua bentuk instrumen yang dapat digunakan yakni tes dan non-tes. Instrumen yang berbentuk tes biasanya untuk mengukur prestasi, seperti prestasi belajar.

Sementara itu, instrumen non-tes pada umumnya digunakan untuk mengukur sikap. Dalam penelitian pendidikan, salah satu instrumen non-tes yang sering digunakan adalah kuisisioner (angket). Secara umum, ada dua jenis kuisisioner yaitu kuisisioner ter-

tutup dan terbuka. Kuisisioner tertutup adalah kuisisioner yang telah disediakan alternatif jawabannya sehingga responden tinggal memilih yang sesuai dengan keadaan dirinya. Sedangkan kuisisioner terbuka adalah kuisisioner yang jawabannya belum disediakan sehingga responden bebas menuliskan apa yang dirasakan. Satu hal yang menjadi ciri utama suatu kuisisioner adalah tidak ada jawaban benar atau salah.

Ketersediaan instrumen aktivitas pengembangan keprofesian berkelanjutan guru sangat diperlukan, mengingat perkembangan terkini menunjukkan penelitian tentang pengembangan keprofesian guru sedang berkembang, baik di lingkungan perguruan tinggi maupun sekolah. Oleh karena itu, penelitian ini berfokus pada pengembangan instrumen asesmen diri terhadap aktivitas pengembangan keprofesian guru secara berkelanjutan. Pengembangan instrumen asesmen diri terhadap aktivitas pengembangan keprofesian berkelanjutan guru sangat penting dalam konteks pengembangan mutu pendidikan. Dengan demikian, penelitian ini diharapkan dapat menghasilkan instrumen asesmen diri terhadap aktivitas pengembangan keprofesian berkelanjutan guru memenuhi validitas dan reliabilitas yang baik. Hasil ini penting untuk ikut membantu meningkatkan ketersediaan instrumen baku yang dibutuhkan para peneliti.

Metode Penelitian

Penelitian ini merupakan penelitian dan pengembangan (*Research and Development*) dengan memadukan model Borg & Gall (1983, p. 775) dengan model pengembangan Plomp. Sepuluh langkah dari model Borg & Gall dimodifikasi menjadi lima tahap (*stage*) seperti tahapan dalam model Plomp (2007, p. 15). Langkah-langkah tersebut adalah: (1) melakukan studi pendahuluan untuk menganalisis produk yang akan dikembangkan, (2) mengembangkan design (3) mengembangkan produk dan validasi produk, (4) melakukan uji coba lapangan, (5) diseminasi/implementasi.

Penelitian dilaksanakan pada bulan Februari sampai April 2015 di provinsi

Nusa Tenggara Barat. Penentuan sampel penelitian dilakukan dengan teknik *purposive random sampling*. Daerah sampel yang dipilih mewakili karakteristik wilayah kota dan kabupaten, yaitu Kota Mataram, Kabupaten Lombok Barat, Kabupaten Lombok Tengah dan Kabupaten Lombok Timur. Distribusi sampel sekolah dalam studi mencerminkan lokasi sekolah, yaitu di ibukota provinsi (SMA di kota Mataram), ibukota kabupaten (SMA di ibukota kabupaten Lombok Barat, Lombok Tengah dan Lombok Timur), dan ibukota kecamatan (SMA di kecamatan dalam kabupaten). Penentuan sampel responden guru (sumber data) sebagai subjek uji, ditetapkan semua guru mata pelajaran di SMA yang terpilih.

Pada tahap uji coba awal dilakukan survei pada sampel kecil, yaitu pada 10 (sepuluh) SMA negeri yang mencerminkan kategori sekolah dengan kualitas baik, sedang dan kurang di wilayah kota/kabupaten berdasarkan rerata nilai ujian nasional. Dalam hal ini, sampel responden pada uji coba tahap awal sebanyak 350 guru. Sementara itu, uji coba tahap akhir dilakukan pada skala yang lebih luas, yaitu pada 60 SMA negeri dan swasta dengan karakteristik yang sama dengan tahap uji coba awal. Dengan demikian, maka jumlah responden sebagai subjek uji pada uji coba tahap akhir adalah sebanyak 950 guru.

Variabel operasional dalam penelitian ini adalah aktivitas pengembangan keprofesian berkelanjutan guru. Adapun spesifikasi isi instrumen dalam penelitian ini dilakukan dengan menjabarkan terlebih dahulu konsep-konsep tentang pengembangan keprofesian berkelanjutan guru menjadi dimensi dan indikator yang mengungkap aktivitas pengembangan keprofesian berkelanjutan guru. Masing-masing indikator dibuat minimal dalam lima butir pernyataan. Jenis respons dalam penelitian ini adalah kinerja tipikal yang tidak dapat dinyatakan benar atau salah, tetapi dapat dikatakan semua respons benar menurut kondisi setiap responden. Sesuai dengan karakteristik jenis respon, maka format alat ukur yang dipilih untuk menyajikan butir-butir instrumen adalah

format pilihan terbatas. Untuk tiap-tiap butir memiliki 5 pilihan jawaban yakni Tidak Pernah (TP), Jarang (JR), Kadang-kadang (KK), Sering (SR), dan Selalu (SL). Penskoran dilakukan dengan memberikan skor 1=tidak pernah, 2=jarang, 3=kadang-kadang, 4=sering, dan 5=selalu.

Analisis validitas konstruk aktivitas pengembangan keprofesian berkelanjutan guru dengan Confirmatory Factor analysis (CFA) dan metode estimasi ML (*Maximum Likelihood*). Sebelum dilakukan analisis *Second Order CFA*, terlebih dahulu dilakukan analisis faktor dengan SPSS 17.0 untuk memastikan terpenuhinya syarat analisis faktor, yaitu kecukupan jumlah data observasi dan adanya korelasi yang signifikan antara variabel-variabel observasi.

Menurut (Hair, Anderson, Tatham, & Black (2010, p. 117), bahwa kecukupan jumlah observasi (data) dapat diidentifikasi melalui parameter *Kaiser-Meyer-Olkin* (KMO) dengan nilai $KMO > 0,5$. Selanjutnya, adanya korelasi antara variabel-variabel multivariat dapat teridentifikasi dengan parameter *Bartlett's Test of Sphericity* yang harus signifikansi dengan *p-value* lebih kecil dari 0,05 ($p\text{-value} < 0,05$). Besarnya korelasi antar-variabel multivariat dapat terlihat dari nilai *Measure of Sampling Adequacy* (MSA) dengan nilai $MSA > 0,5$.

Analisis validitas konstruk model pengukuran dengan CFA pada *Second Order CFA* bertujuan untuk menguji kecocokan model yang diteorikan dengan data empirik. Kriteria utama kecocokan model dengan data lapangan jika terpenuhi minimal tiga syarat yaitu $\text{Chi-square } (p\text{-value}) > 0,05$ dan $\text{Root Mean Square Error of Approximation } (RMSEA) < 0,08$ serta $\text{Comparative Goodness of Fit Indeks } (CFI) \geq 0,9$. Jika model cocok dengan data lapangan, maka analisis dilakukan terhadap muatan faktor setiap indikator atau butir dalam model pengukuran. Namun, apabila sebaliknya, yaitu model tidak cocok dengan data lapangan analisis terhadap muatan faktor tidak dapat dilakukan. Langkah yang harus dilakukan adalah melakukan modifikasi model dengan mempertimbangkan indeks modifikasi pada output Lisrel.

Penentuan validitas butir instrumen dengan *second order* CFA didasarkan pada pendapat Hair et al. (2010, p. 119) bahwa kriteria minimal nilai muatan faktor (λ) yang dapat diterima adalah $\pm 0,3$ sampai $0,4$ dan nilai- $t > 1,96$ dengan taraf signifikansi 5%. Dengan demikian, maka dalam penelitian ini digunakan kriteria nilai muatan faktor, (λ) $\geq 0,3$ dan nilai- $t > 1,96$ untuk masing-masing butir instrumen. Selain itu, nilai- t muatan faktor tidak boleh negatif. Apabila kedua persyaratan muatan faktor sebuah butir tidak terpenuhi, maka butir dimaksud harus dihilangkan dari model serta dilakukan analisis ulang. Kemudian reliabilitas konstruk dianalisis dengan menggunakan formula *Construct Reliability* dengan pertimbangan bahwa konstruk instrumen aktivitas guru cenderung mengarah pada multidimensi (Widhiarso & Mardapi, 2010, p. 10). Kriteria minimal nilai reliabilitas konstruk (CR) menurut Ghozali & Fuad (2012, p. 268) adalah $CR \geq 0,7$.

Hasil Penelitian dan Pembahasan

Data hasil penelitian pada tahap pendahuluan diperoleh melalui kegiatan pengumpulan data dan informasi lapangan serta studi pendahuluan. Kegiatan dimaksud meliputi: observasi awal di berbagai SMA untuk identifikasi masalah riil di lapangan, kajian pustaka, kajian peraturan, dan kajian hasil penelitian terdahulu. Tahapan ini bertujuan untuk menggali informasi aktual tentang implementasi pengembangan keprofesian berkelanjutan guru di sekolah.

Hasil studi pendahuluan dengan observasi awal di berbagai SMA menunjukkan, bahwa pengembangan keprofesian berkelanjutan guru dilaksanakan secara mandiri dan kelompok. Pada umumnya aktivitas pengembangan keprofesian guru meliputi, (1) membaca melalui media cetak dan elektronik; (2) mengembangkan perangkat pembelajaran dan penilaian; (3) pendidikan dan pelatihan fungsional dan kolektif, antara lain IHT, workshop, diklat, MGMP dan seminar; (4) membuat karya tulis berupa modul dan penelitian tindakan kelas (PTK); (5) refleksi proses pembelajaran yang dilakukan

dengan peserta didik dan teman sejawat; dan (6) bekerjasama dengan teman sejawat dengan cara berbagi informasi dan bahan tentang konten terkini yang terkait dengan kurikulum, pembelajaran dan penilaian. Namun tidak sedikit guru mengeluhkan minimnya kesempatan untuk mengikuti pelatihan yang difasilitasi oleh pemerintah.

Disamping itu, aktivitas pengembangan keprofesian tersebut dilaksanakan secara insidental dan mandiri serta kurang terencana dengan baik, walaupun di sekolah telah memiliki unit kerja yang berfungsi untuk mengkordinasikan pengembangan keprofesian berkelanjutan bagi guru.

Berdasarkan kajian terhadap berbagai literatur tentang pengembangan keprofesian berkelanjutan guru, dirumuskan tujuh dimensi menjadi konstruk teoritik instrumen aktivitas pengembangan keprofesian berkelanjutan guru, yaitu: (1) aktivitas membaca, (2) pengembangan pembelajaran dan penilaian, (3) pendidikan dan pelatihan, (4) publikasi ilmiah, (5) publikasi karya inovatif, (6) reflektif, dan (7) aktivitas kolaboratif.

Hasil validasi ahli menunjukkan, bahwa para ahli memberikan penilaian yang relatif sama terkait konstruk aktivitas pengembangan keprofesian berkelanjutan guru. Secara umum indikator yang disusun dinilai sudah mewakili dimensi dari konstruk aktivitas pengembangan keprofesian berkelanjutan guru. Indikator-indikator yang disusun merupakan representasi dari penjabaran dimensi dari konstruk yang telah didefinisikan. Dengan demikian, konstruksi butir sudah sesuai dengan indikatornya.

Para ahli memberikan beberapa masukan terhadap butir-butir instrumen aktivitas pengembangan keprofesian yang disusun berdasarkan indikatornya. Ada beberapa butir instrumen yang saling tumpang tindih (*overlap*) serta mirip dalam satu dimensi sehingga butir tersebut harus direduksi. Hasil kajian para ahli terhadap perangkat instrumen menunjukkan bahwa secara umum instrumen tersebut cukup memadai jika dilihat dari penggunaan tata bahasa dan penulisannya. Bahasa yang digunakan dalam penyusunan perangkat in-

strumen dinilai oleh para ahli cukup komunikatif dan mudah dimengerti. Selanjutnya, para pakar memberikan beberapa masukan untuk penyempurnaan instrumen, antara lain agar melengkapi indikator dan kriteria penskoran supaya lebih jelas, lebih operasional, mudah dipahami serta tidak membingungkan.

Uji keterbacaan instrumen melibatkan praktisi dari unsur-unsur yang terkait dengan calon pengguna model asesmen, yaitu guru, kepala sekolah, pengawas dan widyaiswara. Penilaian terhadap keterbacaan instrumen diarahkan pada aspek: (1) kejelasan petunjuk instrumen, (2) kejelasan indikator aktivitas membaca, (3) kejelasan indikator pengembangan pembelajaran dan penilaian, (4) kejelasan indikator pendidikan dan pelatihan, (5) kejelasan indikator publikasi ilmiah, (6) kejelasan indikator publikasi karya ilmiah, (7) kejelasan indikator aktivitas reflektif, (8) kejelasan indikator aktivitas kolaboratif, (9) penggunaan bahasa Indonesia baku, (10) rumusan pernyataan yang komunikatif, (11) penggunaan kalimat dan kata yang mudah dipahami, (12) tata tulis diarahkan pada penilaian bentuk huruf, ukuran huruf, dan format atau *lay out* instrumen. Adapun hasil penilaian dari dua tahap uji coba terhadap keterbacaan instrumen disajikan pada Tabel 1.

Penilaian keterbacaan instrumen menggunakan skala 5, yaitu skor minimal 1 dan skor maksimal 5. Tabel 1 menunjukkan rerata skor total pada uji coba ke-2 adalah 3,86. Apabila dikonsultasikan dengan pedoman konversi data kuantitatif ke kualitatif, berada pada rerata skor > 3,4 – 4,2 dan termasuk dalam klasifikasi layak atau baik.

Hal ini memberi indikasi bahwa tingkat keterbacaan instrumen yang dikembangkan ini dapat diklasifikasikan layak atau baik.

Uji Coba Tahap Awal

Model hipotetik yang diuji secara empiris dalam penelitian ini adalah konsrak instrumen asesmen diri guru terhadap aktivitas pengembangan keprofesian berkelanjutan dengan tujuh dimensi, yaitu: (1) aktivitas membaca, (2) pengembangan pembelajaran dan penilaian, (3) pendidikan dan pelatihan, (4) publikasi ilmiah, (5) publikasi karya ilmiah, (6) aktivitas reflektif, (7) aktivitas kolaboratif. Model instrumen yang dikembangkan dinyatakan cocok dengan data lapangan apabila sudah terpenuhi dua kriteria dari tiga kriteria yang menjadi ukuran kecocokan yaitu *Root Mean Square Error of Approximation* (RMSEA) < 0,08; Chi-Square yang diperoleh dari pengujian memiliki probabilitas lebih besar dari 0,05 ($p > 0,05$); dan *Comparative Goodness of Fit Indeks* (CFI) $\geq 0,90$.

Hasil analisis faktor dengan program SPSS 17.0 menunjukkan bahwa nilai KMO sebesar 0,921 (KMO > 0,50). Sementara itu, nilai Bartlett's Test of Sphericity menunjukkan signifikansi *p-value* lebih kecil dari 0,05 ($p\text{-value} < 0,05$). Hasil ini menunjukkan telah terpenuhinya kecukupan jumlah observasi (data) dan adanya korelasi yang signifikan antara variabel-variabel observasi aktivitas pengembangan keprofesian berkelanjutan guru. Dengan demikian, maka data hasil observasi telah memenuhi syarat untuk analisis faktor konfirmatori.

Tabel 1. Rekapitulasi Hasil Uji Keterbacaan Instrumen Asesmen Diri pada Aktivitas PKB Guru SMA

Tahap	Jml. Resp.	Rerata Skor												Rerata Total
		1	2	3	4	5	6	7	8	9	10	11	12	
1	20	4,00	3,74	3,63	3,79	3,68	3,79	3,63	3,68	3,63	3,74	3,79	3,79	3,78
2	40	4,03	3,82	3,82	3,87	3,87	3,87	3,79	3,74	3,85	3,74	3,95	3,82	3,86

Nilai *Measure of Sampling Adequacy* (MSA) menunjukkan besarnya korelasi antar variabel multivariat. Menurut Hair (2010, p 120), bahwa batas nilai yang dapat diterima $MSA > 0,50$. Hasil analisis terhadap nilai MSA menunjukkan terdapat 14 dari 55 butir memiliki variansi $MSA < 0,5$ sehingga 14 butir tersebut dieleminasi dan tidak disertakan pada tahapan analisa berikutnya. Selanjutnya, jika dilihat kedudukan butir untuk setiap faktornya setelah dilakukan rotasi nampak bahwa semua butir dapat menjelaskan kedudukan butir dalam faktornya.

Dimensionalitas alat ukur dapat dilihat dari parameter nilai eigen pada setiap faktor. Hasil analisis terhadap variabel aktivitas guru menunjukkan parameter nilai eigen sebesar 13,07 dan variansi 31,12 % dengan jumlah faktor maksimum yang terbentuk 7 faktor, yaitu membaca, mengembangkan pembelajaran dan penilaian, diklat, publikasi ilmiah, publikasi karya inovatif, refleksi, dan kolaborasi. Hal ini menunjukkan, bahwa variansi masing-masing variabel > 20 % sehingga dapat disimpulkan bahwa masing-masing variabel asesmen aktivitas PKB guru berdimensi satu (unidimensi).

Namun demikian, uji kecocokan model secara keseluruhan menunjukkan bahwa kontrak model pengukuran memiliki kecocokan model yang kurang baik (model tidak fit) dengan data lapangan, karena nilai χ^2 (p -value) = $0,00 < 0,05$; $CFI \geq 0,98$ dan $RMSEA < 0,047$. Selanjutnya, langkah untuk memperoleh model yang fit adalah melakukan modifikasi terhadap model. Ghazali & Fuad (2012, p. 170) menyarankan dua tipe modifikasi yang dapat dilakukan untuk memperoleh fit model yang baik, yaitu: (1) dengan menambah path (jalur) dari faktor (subdimensi) ke indikator yang berdasarkan analisis rasional merupakan cerminan faktor tersebut; dan (2) dengan mengkorelasikan antara dua *error indicator* (butir) yang memiliki residual kovarians yang besar.

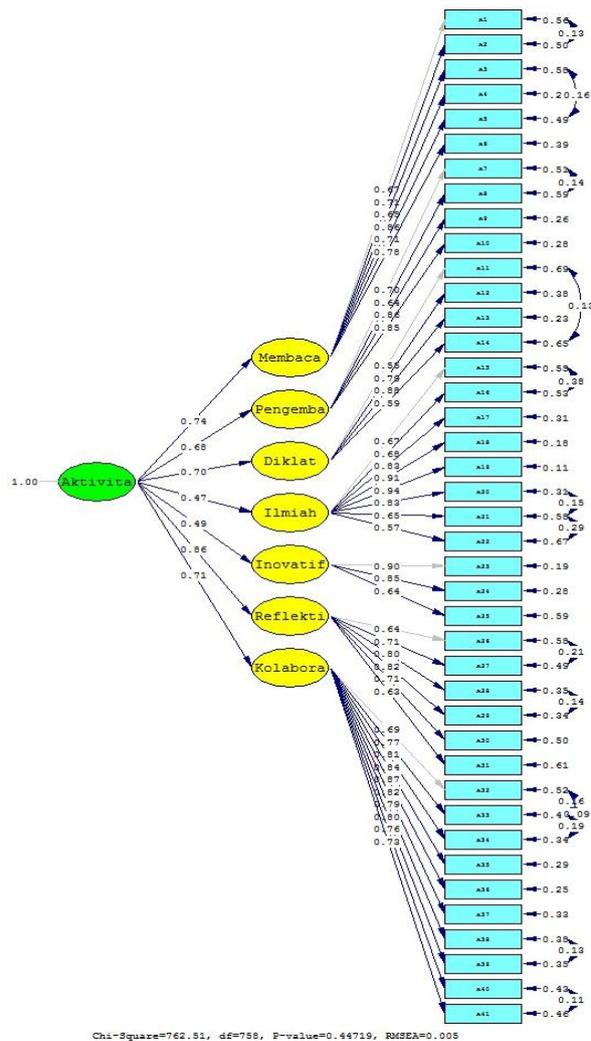
Modifikasi model yang dipilih adalah dengan mengkorelasikan antara dua *error indicator* yang memiliki *residual covarians* yang besar, yaitu: (1) butir A3-A5 dalam dimensi membaca, (2) butir A7-A8 dalam dimensi

pengembangan pembelajaran dan penilaian, (3) butir A15-A16 dan A21-A22 pada dimensi publikasi karya ilmiah, (4) butir A28-A29 pada dimensi aktivitas reflektif, dan (5) butir A32-A34 dan A33-A34 pada dimensi aktivitas kolaboratif. Alternatif ini ditempuh karena tipe modifikasi ini bersifat tidak mengubah kontrak sebagaimana yang telah dihipotesiskan. Selain itu, secara konseptual korelasi antara indikator-indikator atau butir tersebut didasari oleh suatu argumen yang kuat bahwa kedua indikator yang dikorelasikan tersebut mengungkap trait yang hampir sama.

Hasil modifikasi terhadap model menunjukkan bahwa secara keseluruhan model pengukuran memiliki kecocokan yang baik (model fit), karena nilai Chi -Square = 817,26 p -value = $0,085 > 0,05$; $RMSEA = 0,018 < 0,08$ dan $CFI = 1,00 > 0,90$. Dengan demikian, maka model pengukuran aktivitas pengembangan keprofesian berkelanjutan guru telah memenuhi kriteria unidimensi, yaitu sebanyak 41 butir (A1 s.d A41) yang merefleksikan dimensi aktivitas. Adapun model pengukuran aktivitas pengembangan keprofesian berkelanjutan guru setelah revisi pada *second order CFA* disajikan pada Gambar 1.

Selain itu, estimasi signifikansi parameter muatan faktor dengan kriteria nilai- $t > 1,96$ dimaksudkan untuk mengetahui kelayakan sebuah butir menjadi alat ukur. Hasil estimasi nilai- t muatan faktor (λ) menunjukkan, bahwa semua butir aktivitas guru memiliki nilai- $t > 1,96$ dengan nilai muatan faktor standar (λ) $> 0,3$ (rentang nilai 0,57 s.d 0,93) serta tidak ada nilai negatif. Selanjutnya, semua dimensi yang dikembangkan juga memiliki koefisien jalur (γ) $> 0,30$ (rentang nilai 0,44 s.d 0,68) dengan nilai- $t > 1,96$. Dengan demikian, dapat disimpulkan bahwa ke-41 butir instrumen terkait dengan aktivitas guru memenuhi syarat dan layak menjadi alat ukur dimensi aktivitas pengembangan keprofesian berkelanjutan guru.

Evaluasi terhadap reliabilitas model pengukuran dilakukan dengan menggunakan *composite reliability measure* (ukuran reliabilitas komposit) atau yang dikenal juga dengan *construct reliability* (reliabilitas kontrak)



Gambar 1. Model Pengukuran Asesmen Diri Guru pada Aktivitas Pengembangan Keprofesian Berkelanjutan

Sebuah kontrak mempunyai reliabilitas yang baik jika nilainya lebih dari atau sama dengan 0,70 (Lance, Butts, & Michels, 2006, p. 205). Berdasarkan hasil perhitungan reliabilitas kontrak, dapat disimpulkan bahwa reliabilitas kontrak model asesmen diri aktivitas pengembangan keprofesian berkelanjutan guru adalah baik ($CR > 0,70$).

Uji Coba Tahap Akhir

Tahapan validasi empirik tahap awal telah menghasilkan kontrak asesmen diri guru pada pengembangan keprofesian berkelanjutan yang direfleksikan oleh tujuh dimensi, yaitu: (1) aktivitas membaca, (2) pengembangan pembelajaran dan penilaian,

(3) pendidikan dan pelatihan, (4) publikasi ilmiah, (5) publikasi karya ilmiah, (6) aktivitas reflektif, (7) aktivitas kolaboratif. Kontrak aktivitas guru pada pengembangan keprofesian berkelanjutan ini tercermin dalam 41 indikator. Tahapan selanjutnya adalah validasi empirik tahap akhir terhadap 41 butir pertanyaan tersebut pada 950 guru SMA.

Hasil analisis faktor dengan program SPSS 17.0 menunjukkan bahwa nilai KMO sebesar 0,941 ($KMO > 0,50$), sedangkan nilai *Bartlett's Test of Sphericity* menunjukkan signifikansi p-value lebih kecil dari 0,05 ($p\text{-value} < 0,05$). Hasil ini menunjukkan telah terpenuhinya kecukupan jumlah observasi (data) dan adanya korelasi yang signifikan antara variabel-variabel observasi aktivitas guru. Dengan demikian, maka data hasil observasi terhadap aktivitas pengembangan keprofesian berkelanjutan guru telah memenuhi syarat analisis faktor konfirmatori.

Hasil validasi empirik tahap akhir terhadap kontrak aktivitas pengembangan keprofesian berkelanjutan guru menunjukkan kecocokan keseluruhan model (*overall model*) yang baik (model fit), karena nilai Chi-Square = 762,51 $p\text{-value} = 0,45 > 0,05$; RMSEA = 0,005 $< 0,08$ dan CFI = 1,00 $> 0,90$. Dengan demikian, maka model pengukuran aktivitas guru telah memenuhi kriteria unidimensi, yaitu sebanyak 41 butir (A1 s.d A41) yang merefleksikan dimensi aktivitas hanya mengukur hal yang sama, yaitu dimensi aktivitas pengembangan keprofesian berkelanjutan guru.

Selain itu, estimasi signifikansi parameter muatan faktor dengan kriteria nilai- $t > 1,96$ dimaksudkan untuk mengetahui kelayakan sebuah butir menjadi alat ukur. Hasil estimasi nilai- t muatan faktor (λ) menunjukkan, bahwa semua butir aktivitas guru memiliki nilai- $t > 1,96$ dengan nilai muatan faktor standar (λ) $> 0,3$ (rentang nilai 0,55 s.d 0,94) serta tidak ada nilai negatif. Selanjutnya, semua dimensi yang dikembangkan juga memiliki koefisien jalur (γ) $> 0,30$ (rentang nilai 0,47 s.d 0,86) dengan nilai- $t > 1,96$. Dengan demikian, dapat disimpulkan bahwa ke-41 butir pertanyaan/

pernyataan yang terkait dengan aktivitas guru memenuhi syarat dan layak menjadi alat ukur aktivitas pengembangan keprofesian berkelanjutan guru.

Evaluasi terhadap reliabilitas model pengukuran dilakukan dengan menggunakan *composite reliability measure* (ukuran reliabilitas komposit) atau yang dikenal juga dengan *construct reliability* (reliabilitas konstruk) (Hair et al., 2010). Sebuah konstruk mempunyai reliabilitas yang baik jika nilainya lebih dari atau sama dengan 0.70 (Hair et al., 2010; Lance et al., 2006, p. 205). Berdasarkan ringkasan hasil perhitungan reliabilitas konstruk, dapat disimpulkan bahwa konstruk model asesmen diri pada aktivitas pengembangan keprofesian berkelanjutan guru adalah reliabel ($CR > 0,70$).

Dari hasil analisis secara kuantitatif dan kualitatif, secara umum dapat dikatakan bahwa instrumen asesmen diri guru terhadap aktivitas pengembangan keprofesian berkelanjutan yang dikembangkan berdasarkan kajian teoretik dan telaah para pakar dan panelis serta uji coba secara empiris sudah sesuai dan dapat diaplikasikan. Jika dibandingkan antara konsep dan draf awal instrumen, terdapat beberapa perbaikan/revisi dan pengembangan sebagai hasil dari validasi oleh pakar dan panelis serta uji coba empiris kepada responden guru di lapangan baik tahap pertama dan tahap kedua. Pengembangan tersebut mencakup kesesuaian antara butir dengan indikator, indikator dengan dimensi pada konstruk latennya, serta penggunaan bahasa yang lebih komunikatif dan mudah dipahami.

Validitas mengukur sejauh mana suatu alat ukur mampu mengukur apa yang seharusnya diukur. Instrumen asesmen ini dikembangkan berdasarkan pada konteks lapangan dan kajian teoritik yang relevan. Hasil penilaian para pakar dan uji coba empiris menunjukkan bahwa instrumen tes ini memiliki validitas isi yang memadai sehingga layak sebagai alat ukur dan dapat digunakan sebagai instrumen pendukung dalam pembinaan dan pengembangan profesi dan karir guru.

Koefisien reliabilitas konstruk uji coba empiris pertama dan kedua lebih dari 0,7. Reliabilitas suatu alat ukur merupakan konsistensi atau keajegan alat ukur tersebut dalam mengukur apa yang diukur. Makin tinggi koefisien reliabilitas, makin dekat nilai skor amatan dengan nilai komponen skor yang sesungguhnya, sehingga nilai skor amatan dapat digunakan sebagai pengganti komponen skor yang sesungguhnya. Dengan demikian dapat dinyatakan bahwa hasil pengukuran dengan menggunakan instrumen ini menunjukkan aktivitas pengembangan keprofesian berkelanjutan guru SMA yang mendekati sesungguhnya.

Simpulan

Simpulan hasil penelitian ini adalah: (1) konstruk asesmen diri guru terhadap aktivitas pengembangan keprofesian berkelanjutan, dicerminkan oleh tujuh dimensi, yaitu: yaitu: (a) aktivitas membaca, (b) pengembangan pembelajaran dan penilaian, (c) pendidikan dan pelatihan, (d) publikasi ilmiah, (e) publikasi karya ilmiah, (f) aktivitas reflektif, dan (g) aktivitas kolaboratif. (2) hasil uji keterbacaan menunjukkan instrumen dalam kategori baik dan layak, rerata skor 3,86; (3) hasil analisis instrumen menunjukkan, bahwa seluruh instrumen valid (muatan faktor $> 0,3$ dan nilai- $t > 1,96$) dan koefisien reliabilitas konstruk $> 0,7$ serta memenuhi syarat sebagai model yang fit (Chi-Square = 762,51 p-value = 0,45 $> 0,05$; RMSEA = 0,005 $< 0,08$ dan CFI = 1,00 $> 0,90$); (4) instrumen asesmen diri yang dikembangkan memenuhi syarat validitas dan reliabilitas sehingga layak digunakan oleh guru sebagai asesmen diri terhadap aktivitas pengembangan keprofesian berkelanjutan.

Berdasarkan simpulan penelitian dapat disampaikan saran kepada peneliti lain di bidang asesmen pendidikan dapat menggunakan produk ini untuk menguji kecocokan model dengan responden guru-guru pada jenjang pendidikan dasar SD/MI dan SMP/MTs karena instrumen asesmen diri guru yang dihasilkan ini hanya divalidasi pada responden guru SMA.

Daftar Referensi

- Bakkenes, I., Vermunt, J. D., & Wubbels, T. (2010). Teacher learning in the context of educational innovation: Learning activities and learning outcomes of experienced teachers. *Learning and Instruction, 20*(6), 533–548. <https://doi.org/10.1016/j.learninstruc.2009.09.001>
- Benson, P. (2007). Autonomy in language teaching and learning. *Language Teaching, 40*(1), 21. <https://doi.org/10.1017/S0261444806003958>
- Borg, W. R., & Gall, M. D. (1983). *Educational research: An introduction* (4th ed.). New York: Longman Publishing.
- Borko, H. (2004). Professional development and teacher learning: mapping the terrain. *Educational Researcher, 33*(8), 3–15. <https://doi.org/10.3102/0013189X033008003>
- Brew, A. (1999). Toward autonomous assessment: using self assessment and peer assessment. In S. Brown & A. Glesner (Eds.), *Assessment Matters in Higher Education*. Open University Press.
- Day, C., & Sachs, J. (2004). *International handbook on the continuing professional development of teachers*. Berkshire: Open University Press.
- de Vries, S., Jansen, E. P. W. A., & van de Grift, W. J. C. M. (2013). Profiling teachers' continuing professional development and the relation with their beliefs about learning and teaching. *Teaching and Teacher Education, 33*, 78–89. <https://doi.org/10.1016/j.tate.2013.02.006>
- Desimone, L. M. (2009). Improving impact studies of teachers' professional development: toward better conceptualizations and measures. *Educational Researcher, 38*(3), 181–199. <https://doi.org/10.3102/0013189X08331140>
- Gaible, E., & Burns, M. (2005). *Using technology to train teachers: Appropriate uses of ICT for teacher professional development in developing countries*. Washington, DC: The International Bank for Reconstruction and Development/World Bank.
- Ghozali, I., & Fuad. (2012). *Structural equation modelling: teori konsep, dan aplikasi dengan Program LISREL 8.80* (3rd ed.). Semarang: Badan Penerbit Universitas Diponegoro.
- Hair, J. F., Anderson, R. E., Tatham, R. L., & Black, W. C. (2010). *Multivariate data analysis* (4th ed.). New Jersey: Prentice Hall.
- Hoekstra, A., Brekelmans, M., Beijaard, D., & Korthagen, F. (2009). Experienced teachers' informal learning: Learning activities and changes in behavior and cognition. *Teaching and Teacher Education, 25*(5), 663–673. <https://doi.org/10.1016/j.tate.2008.12.007>
- Kelchtermans, G. (2004). CPD for professional renewal: moving beyond knowledge for practice dalam . Open University Press McGraw-Hill Education. In C. Day & J. Sachs (Eds.), *International Handbook on the Continuing Professional Development of Teachers* (pp. 217–237). Berkshire: Open University Press.
- Kwakman, K. (2003). Factors affecting teachers' participation in professional learning activities. *Teaching and Teacher Education, 19*(2), 149–170. [https://doi.org/10.1016/S0742-051X\(02\)00101-4](https://doi.org/10.1016/S0742-051X(02)00101-4)
- Lance, C. E., Butts, M. M., & Michels, L. C. (2006). The sources of four commonly reported cutoff criteria. *Organizational Research Methods, 9*(2), 202–220. <https://doi.org/10.1177/1094428105284919>

- Menteri Pendidikan dan Kebudayaan Republik Indonesia. (2010). *Buku I: Pedoman pengelolaan pengembangan keprofesionalan berkelanjutan*. Jakarta: Kemdikbud.
- Mullis, I. V. S., Martin, M. O., Foy, P., & Arora, A. (2012). *TIMSS 2011 international results in mathematics*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
- Plomp, T. J. (2007). Educational design research: an introduction. In *An Introduction to Educational Research*. Enschede: National Institute for Curriculum Development.
- Rolheiser, C., & Ross, J. A. (2005). Students Self-evaluation: what research says and what practice shows. Retrieved from [http://moodle.manistee.org/pluginfile.php/59439/course/section/16807/STUDENT_SELF-EVALUATION WHAT RESEARCH SAYS AND WHAT PRACTICE SHOWS.pdf](http://moodle.manistee.org/pluginfile.php/59439/course/section/16807/STUDENT_SELF-EVALUATION_WHAT_RESEARCH_SAYS_AND_WHAT_PRACTICE_SHOWS.pdf)
- Tynjälä, P. (2008). Perspectives into learning at the workplace. *Educational Research Review*, 3(2), 130–154. <https://doi.org/10.1016/j.edurev.2007.12.001>
- Widhiarso, W., & Mardapi, D. (2010). Komparasi ketepatan estimasi koefisien reliabilitas teori skor murni klasik. *Jurnal Penelitian Dan Evaluasi Pendidikan*, 14(1). Retrieved from <http://journal.uny.ac.id/index.php/jpep/article/view/1973>
- Wurinanda, I. (2015, December 30). Rata-rata nilai UKG di bawah standar. *Okezone News*. Retrieved from <https://news.okezone.com/read/2015/12/30/65/1277618/rata-rata-nilai-ukg-di-bawah-standar>
- Zubaidah, N. (2012, August 3). Hasil uji kompetensi guru memprihatinkan. *Sindonews.com*. Retrieved from <https://nasional.sindonews.com/read/662873/15/hasil-uji-kompetensi-guru-memprihatinkan-1344004590>

PENGEMBANGAN MODEL AKREDITASI SEKOLAH MENENGAH ATAS/MADRASAH ALIYAH (SMA/MA)

Marjuki ^{1*}, Djemari Mardapi ², Badrun Kartowagiran ²

¹LPMP Bengkulu, ²Universitas Negeri Yogyakarta

¹Jl. Zainul Arifin, Tim. Indah, Singaran Pati, Kota Bengkulu, Bengkulu 38225, Indonesia

²Jl. Colombo No. 1, Depok, Sleman 55281, Yogyakarta, Indonesia

* Corresponding Author. Email: marjuki.pep@gmail.com

Abstrak

Penelitian ini bertujuan memperoleh model akreditasi SMA/MA yang efektif, efisien, dan akuntabel. Produk model yang dikembangkan adalah instrumen akreditasi dan prosedur akreditasi. Penelitian ini termasuk jenis penelitian pengembangan. Pengembangan model dilakukan melalui kegiatan *focus group discussion* (FGD). Validasi instrumen dilakukan oleh pakar, sedangkan validasi prosedur dilakukan oleh peserta FGD. Validitas isi instrumen ditentukan dengan cara menghitung indeks koefisien V Aiken, sedangkan estimasi reliabilitas instrumen ditentukan dengan teknik *intraclass correlation coefficient* (ICC) menggunakan SPSS. Validasi prosedur dilakukan dengan cara menghitung rerata skor hasil penilaian peserta FGD. Uji coba dilakukan untuk mengetahui keefektifan model yang dikembangkan. Subjek coba adalah praktisi terdiri dari kepala sekolah, guru, dan tenaga administrasi. Keefektifan model ditentukan dengan cara menghitung rerata skor hasil penilaian praktisi. Model akreditasi yang dikembangkan diberi nama Model Akreditasi SMART. Kesimpulan tentang Model Akreditasi SMART: (1) memiliki karakteristik sebagai model akreditasi SMA/MA yang efektif, efisien, dan akuntabel; (2) instrumen yang digunakan memiliki validitas dalam kategori sangat baik dan reliabilitas dalam kategori baik; (3) prosedur akreditasi seperti dipaparkan di atas mendapat penilaian baik dari peserta FGD; dan (4) keefektifan model termasuk kategori baik dan dapat digunakan.

Kata kunci: *model akreditasi, instrumen akreditasi, prosedur akreditasi*

DEVELOPING AN ACCREDITATION MODEL OF SENIOR SECONDARY SCHOOL (SMA/MA)

Abstract

This study is aimed at obtaining an effective, efficient and accountable accreditation model for Senior Secondary School/Madrasah Aliyah (SMA/MA). The product of the model developed is an accreditation instrument and procedure. This study is a research and development (R & D). The model developing was conducted through focus group discussion (FGD). Instrument validation was done by experts, while the procedure validation was performed by the FGD participants. The content validity of the instrument was determined by calculating the V Aiken coefficient index, while the instrument reliability estimation was determined by using intraclass correlation coefficient (ICC) using SPSS. Procedure validation was done by calculating the mean score of the FGD participants' assessment results. Trial was conducted to find out the effectiveness of the developed model. The subject of the trial was practitioners comprising of principals, teachers, and administrative personnels. The effectiveness of the model was determined by calculating the average score of the practitioners' assessment. The accreditation model developed was named SMART Accreditation Model. The conclusion on the SMART Accreditation Model are as follows: (1) having the characteristics of an effective, efficient, and accountable accreditation model for SMA/MA; (2) the content validity and reliability of the instrument used are in a very good category; (3) the afore-mentioned accreditation procedure is considered to be good by the FGD participants; and (4) the effectiveness of the model is in a good category and it can be used.

Keywords: *accreditation model, accreditation instrument, accreditation procedure*

Permalink/DOI: <http://dx.doi.org/10.21831/jpep.v22i1.9860>

Pendahuluan

Tolok ukur atau parameter mutu pendidikan pada jenjang pendidikan dasar dan menengah adalah standar nasional pendidikan (SNP). SNP merupakan kriteria minimal tentang sistem pendidikan di seluruh wilayah hukum Negara Kesatuan Republik Indonesia seperti yang telah dijelaskan dalam Peraturan Pemerintah Republik Indonesia Nomor 19 Tahun 2005 tentang Standar Nasional Pendidikan (Menteri Pendidikan Nasional, 2005). SNP meliputi: (1) standar isi, (2) standar proses, (3) standar kompetensi lulusan, (4) standar pendidik dan tenaga kependidikan, (5) standar sarana dan prasarana, (6) standar pengelolaan, (7) standar pembiayaan, dan (8) standar penilaian pendidikan.

Dalam rangka meningkatkan mutu pendidikan nasional, pemerintah telah melakukan berbagai upaya seperti halnya pengembangan dan penyempurnaan kurikulum, pengembangan materi pembelajaran, perbaikan sistem evaluasi, pengadaan buku dan alat-alat pelajaran, perbaikan sarana prasarana pendidikan, peningkatan kompetensi guru, serta peningkatan mutu pimpinan sekolah (Depdiknas, 2002, p. 3). Namun demikian, upaya tersebut sampai sekarang belum menunjukkan hasil sebagaimana yang diharapkan.

Satuan pendidikan (sekolah/madrasah) dikatakan bermutu apabila satuan pendidikan tersebut telah memenuhi kriteria minimal dari setiap komponen SNP. Fokus peningkatan mutu pendidikan harus diarahkan pada bagaimana mewujudkan suasana belajar dan proses pembelajaran agar peserta didik dapat secara aktif mengembangkan potensi dirinya. Dengan berkembangnya potensi peserta didik secara optimal, maka berdampak pada pencapaian atau pemenuhan kriteria minimal pada komponen standar kompetensi lulusan. Menurut Mardapi (2003, p. 8), usaha peningkatan mutu pendidikan (tercapainya standar mutu kompetensi lulusan) dapat ditempuh melalui peningkatan mutu pembelajaran dan mutu sistem penilaian. Dengan kata lain, meningkatnya mutu pembelajaran yang dilaksana-

kan di tingkat satuan pendidikan akan meningkatkan mutu kompetensi lulusan.

Agar terjadi peningkatan mutu kompetensi lulusan, perlu dilakukan upaya yang sistematis dan terpadu melalui peningkatan komponen standar pendidikan nasional yang lain. Kegiatan sistemik dan terpadu dalam upaya pemenuhan standar-standar mutu pendidikan disebut penjaminan mutu pendidikan. Penjaminan mutu pendidikan oleh satuan pendidikan sesuai dengan Peraturan Mendiknas Nomor 63 Tahun 2009 tentang Sistem Penjaminan Mutu Pendidikan ditujukan untuk memenuhi tiga tingkatan acuan mutu, yaitu: (1) standar pelayanan minimal, (2) standar nasional pendidikan, dan (3) standar mutu pendidikan di atas SNP (Menteri Pendidikan Nasional, 2009). Penjaminan dan pengendalian mutu pendidikan yang sesuai dengan standar nasional pendidikan dilakukan melalui evaluasi, akreditasi, dan sertifikasi.

Akreditasi sekolah bertujuan memberikan pengakuan atas capaian standar mutu pendidikan yang ditetapkan dengan pemberian sertifikat status akreditasi. Objek penilaian dalam akreditasi sekolah mencakup delapan komponen standar nasional pendidikan. Akreditasi sekolah adalah proses penilaian secara komprehensif terhadap satuan atau program pendidikan (Menteri Pendidikan Nasional, 2007, p. 1).

Dalam *Accreditation Handbook* (COA, 2008, p. 3), akreditasi didefinisikan sebagai proses verifikasi yang dilakukan oleh suatu lembaga independen terhadap program melalui penilaian secara menyeluruh apakah menunjukkan telah memenuhi atau melebihi standar umum dan standar program yang ditetapkan. Program yang akreditasi yang dinilai efektif dalam mempersiapkan pendidik dan menunjukkan mutu secara keseluruhan terhadap program-program umum yang ditetapkan. Status akreditasi dapat dicapai meskipun tidak semua standar dipenuhi.

Pagliarulo (1996, pp. 1114–1115) mengemukakan tiga aspek dalam akreditasi: (1) proses akreditasi diselenggarakan oleh suatu organisasi atau lembaga eksternal, (2) unit

yang diakreditasi dapat berupa program studi atau satuan pendidikan, dan (3) standar-standar tertentu harus ditentukan terlebih dulu.

ECA (2005, p. 4) mendefinisikan akreditasi sebagai suatu keputusan mandiri dan formal yang diperoleh dari suatu lembaga mandiri berdasarkan standar tertentu. Proses akreditasi dilaksanakan melalui beberapa tahapan mulai dari penilaian diri (*self-assessment*), penyiapan berbagai dokumen yang perlu diajukan menjalani akreditasi, penilaian eksternal oleh tim asesmen yang independen, dan penetapan keputusan akreditasi. Keputusan akreditasi tergantung pada penilaian mutu berdasarkan standar mutu yang telah ditetapkan. Keputusan akhir dari sebuah proses atau prosedur akreditasi yang berwibawa perlu dilakukan secara wajar meskipun diakui bahwa validitas terhadap hasil penilaian itu masih terbatas.

Dalam Undang-Undang Republik Indonesia Nomor 20 Tahun 2003 tentang Sistem Pendidikan Nasional (Presiden Republik Indonesia, 2003), disebutkan bahwa akreditasi dilakukan untuk menentukan kelayakan program dan satuan pendidikan pada jalur pendidikan formal dan nonformal pada setiap jenjang dan jenis pendidikan. Proses akreditasi dilakukan secara berkala dan terbuka dengan tujuan untuk membantu dan memberdayakan program dan satuan pendidikan agar mampu mengembangkan sumber daya yang ada dalam rangka mencapai tujuan pendidikan nasional.

Dalam proses akreditasi terdapat kegiatan pengukuran, penilaian, dan evaluasi. Terkait dengan istilah ini, Huitt (2004, p. 1) memberikan konsep yang berbeda antara pengukuran, penilaian, dan evaluasi. Pengukuran adalah proses kuantifikasi data. Penilaian adalah proses pengumpulan data untuk memahami sebuah isu atau objek secara lebih baik. Evaluasi adalah proses membandingkan antara data yang ada dengan standar atau kriteria yang telah ditetapkan. Hasil evaluasi selanjutnya digunakan sebagai pertimbangan dalam pengambilan keputusan.

Menurut Mardapi (2008, p. 1), penilaian mencakup semua cara yang digunakan

untuk menilai unjuk kerja individu atau kelompok. Proses penilaian meliputi pengumpulan bukti-bukti tentang pencapaian suatu program. Dalam konteks ini, penilaian dapat berkaitan dengan semua proses pendidikan, antara lain: kurikulum, sarana dan prasarana, administrasi, metode pembelajaran, dan peserta didik. Selain itu, Mardapi (2008, p.8) juga mengemukakan bahwa evaluasi merupakan salah satu rangkaian kegiatan dalam meningkatkan mutu, kinerja, atau produktivitas suatu lembaga dalam melaksanakan programnya. Semua program kegiatan menggunakan evaluasi untuk melihat tingkat keberhasilan program yang telah dicapai, mengetahui Keefektifan dan efisiensi program yang sedang berjalan dan memperoleh informasi untuk penetapan kegiatan berikutnya.

Griffin & Nix (1991, p. 15) berpendapat bahwa pengukuran, penilaian, dan evaluasi merupakan sebuah hirarki. Pengukuran adalah kegiatan membandingkan antara hasil observasi dengan kriteria yang telah ditetapkan. Penilaian adalah kegiatan menginterpretasi dan mendeskripsikan hasil pengukuran berdasarkan bukti-bukti yang ada. Evaluasi adalah *judgment* atas implikasi perilaku hasil penilaian.

Bedasarkan konsep pengukuran, penilaian, dan evaluasi sebagaimana diuraikan di atas, dapat dikatakan bahwa dalam proses akreditasi sekolah terdapat kegiatan pengukuran, penilaian, dan evaluasi. Pengukuran terhadap objek yang dinilai menggunakan instrumen. Penilaian dilakukan melalui pengumpulan dan interpretasi data hasil pengukuran disesuaikan dengan kondisi nyata berdasarkan bukti-bukti nyata yang ada di sekolah. Selanjutnya evaluasi adalah proses pengambilan keputusan atau *judgment* dengan menggunakan acuan data hasil penilaian dan pengukuran, berdasarkan standar atau kriteria yang telah ditetapkan. Hasil keputusan berupa penetapan status akreditasi sekolah.

Akreditasi sekolah harus dimaknai sebagai upaya meningkatkan mutu, kinerja, dan produktivitas satuan pendidikan. Ada tiga hal penting dalam proses akreditasi

sekolah agar hasilnya dapat meningkatkan mutu pendidikan, yaitu: masukan, proses, dan keluarannya. Akreditasi merupakan salah satu bentuk evaluasi yang penting dalam rangka mempertahankan mutu pendidikan (Patil & Pudlowski, 2005, pp. 49–58). Ada tiga tahapan proses pendidikan, yaitu: masukan (*input*), proses (*process*), dan hasil (*output*). Oleh karena itu, dalam proses akreditasi perlu dilakukan penilaian secara menyeluruh mencakup aspek-aspek input, proses, dan hasil.

Hasil akreditasi sekolah juga dapat digunakan sebagai bahan evaluasi. Tujuan evaluasi adalah untuk melihat apakah program telah direncanakan tercapai atau belum, berharga atau tidak, dan apakah efisien dalam pelaksanaannya. Evaluasi mutu satuan pendidikan merupakan hal penting yang harus dilakukan oleh pihak satuan pendidikan itu sendiri, pihak penyelenggara pendidikan, dan pihak eksternal.

Kartowagiran (2006, p. 1), mengatakan evaluasi adalah suatu metode untuk mengetahui dan menilai keefektifan suatu program dengan membandingkan kriteria yang telah ditentukan atau tujuan yang ingin dicapai dengan hasil yang dicapai. Hasil yang dicapai dalam bentuk informasi, digunakan sebagai bahan pertimbangan untuk pembuatan keputusan dan penentuan kebijakan.

Joint Committee on Standard Evaluation dalam Stufflebeam & Shinkfield (2007, p. 3), menyebutkan evaluasi adalah penilaian sistematis dari suatu objek yang bernilai atau memiliki kebaikan. Pendapat lain dikemukakan oleh Stark & Thomas (1994, p. 12) yang menyatakan bahwa evaluasi adalah proses untuk memastikan suatu keputusan, memilih informasi yang tepat, mengumpulkan dan menganalisis informasi dalam rangka memberikan laporan data secara singkat yang berguna dalam memilih diantara berbagai alternatif untuk pengambilan keputusan.

Secara formal ada enam fungsi evaluasi pendidikan sebagaimana dikemukakan oleh (Worthen & Sander (1989, p. 21), yaitu: (1) menyiapkan dasar untuk membuat keputusan dan kebijakan, (2) melakukan penilaian hasil belajar, (3) melakukan evaluasi

kurikulum, (4) melaksanakan akreditasi sekolah, (5) monitoring penggunaan dana masyarakat, dan (6) memperbaiki materi dan program pendidikan. Tujuan evaluasi pendidikan adalah upaya meningkatkan mutu pendidikan.

Model evaluasi yang lazim diterapkan untuk keperluan akreditasi sekolah adalah model evaluasi *professional judgment*. Asumsi dasar yang digunakan adalah “*Best evaluation is the expert opinion of qualified professional*” (Artcraft, 2005, p. 133). Evaluasi terbaik adalah pendapat ahli profesional yang memenuhi syarat. Evaluasi model *personal judgment* ini memfokuskan perhatian pada proses pendidikan dengan menggunakan pertimbangan profesional dan pengembangan standar-standar untuk program pendidikan. Model evaluasi *personal judgment* dapat disebut model SAVE yaitu: *Self-study*, *Visitation*, *Annual report*, dan *Evaluation panel*.

Melakukan penilaian mutu sekolah sama halnya dengan melakukan evaluasi mutu sekolah yang harus dilakukan dengan prinsip-prinsip dan standar-standar tertentu. *The Joint Committee on Standards for Educational Evaluation* (Stufflebeam & Shinkfield, 2007, pp. 90-92) memberikan standar-standar dalam melaksanakan evaluasi program sekolah yaitu: manfaat (*utility*), kelayakan (*feasibility*), kesopanan (*propriety*), dan ketepatan (*accuracy*).

Evaluasi mutu sekolah harus dilakukan secara komprehensif yakni menyangkut semua aspek yang terlibat dalam upaya mewujudkan tujuan pendidikan yang telah ditetapkan. Thorndike, Cunningham, Thorndike, & Hagen (1991, p. 58) menyatakan bahwa tujuan dan kegunaan evaluasi pada pendidikan diarahkan kepada keputusan-keputusan yang menyangkut: (a) pengajaran, (b) hasil belajar, (c) diagnosis dan usaha-usaha perbaikan, (d) penempatan, (e) seleksi, (f) bimbingan dan konseling, (g) kurikulum, dan (h) penilaian kelembangaan. Agar keputusan yang diambil melalui evaluasi merupakan keputusan yang tepat, maka evaluasi mutu satuan pendidikan harus dirumuskan berdasarkan komponen-komponen dan indikator-indikator yang tepat.

Sutadji (2009, p. 54) menyatakan bahwa evaluasi terhadap mutu satuan pendidikan memerlukan kajian yang menyangkut komponen-komponen, indikator-indikator, dan kriteria yang digunakan. Acuan yang dapat digunakan untuk menilai mutu satuan pendidikan antara lain model sekolah efektif, model sekolah bermutu, dan model determinan mutu pendidikan.

Beberapa penelitian yang relevan dengan akreditasi sekolah dapat dipaparkan sebagai berikut. Penelitian Lindsay & Campbell (2002, pp. 29–31) tentang pengujian status akreditasi sebagai salah satu indikator mutu program pendidikan menyatakan bahwa status akreditasi sering dianggap sebagai indikator mutu program pendidikan. Penelitian Bennett (2000, pp. 1–6) tentang akreditasi sekolah independen sebagai pedoman untuk memilih sekolah, diperoleh informasi bahwa akreditasi terhadap program pendidikan di sekolah dilakukan untuk mengetahui apakah isi program pendidikan yang diselenggarakan sekolah, apakah sekolah memiliki standar-standar yang tinggi, dan apakah ada komitmen dari sekolah untuk memperbaiki dan meningkatkan mutu programnya. Penelitian Sumarno (2000, p. 58) tentang pengembangan model akreditasi SLTP dan SMU merekomendasikan antara lain: model akreditasi sekolah harus memenuhi prinsip-prinsip pengukuran dan evaluasi dengan subjek satuan analisis sekolah dan objek pokok dari akreditasi sekolah adalah mutu sekolah, kemampuan, dan kinerja sekolah dalam penyelenggaraan pendidikan yang bermutu.

Berdasarkan hasil diskusi sesama anggota BAP-S/M provinsi Bengkulu, asesor akreditasi SMA/MA, dan pengalaman sebagai asesor akreditasi SMA/MA, dapat dikemukakan beberapa permasalahan dalam pelaksanaan dan hasil akreditasi, antara lain: (a) belum ditegakkan sepenuhnya prinsip-prinsip akreditasi dalam proses penilaian kelayakan sekolah/madrasah, (b) masih ada asesor yang belum memahami standar nasional pendidikan dan perangkat akreditasi yang digunakan, (c) masih ditemukan asesor

yang kurang teliti saat melakukan verifikasi dan validasi data.

Aswandi (2009, p. 1) mengemukakan bahwa: (a) belum adanya penilaian kinerja bagi asesor, (b) akreditasi belum berjalan secara efektif, (c) rekomendasi hasil akreditasi belum digunakan sebagai bahan pertimbangan dalam pembinaan sekolah, (d) kurang tersedianya waktu yang cukup untuk melaksanakan akreditasi terutama dalam memilih waktu yang tepat dengan memperhatikan kondisi alam dan geografis dan hari libur sekolah.

Hasil survei keterlaksanaan akreditasi sekolah/madrasah yang dilakukan peneliti pada tahun 2009 dengan melibatkan responden sebanyak 45 orang terdiri dari kepala sekolah/madrasah, guru, dan tenaga administrasi dari 5 SMA/MA di Kota Yogyakarta diperoleh informasi bahwa: (1) sebanyak 37,78% responden mengaku sekolahnya kurang siap diakreditasi; (2) sebanyak 24,44% responden menyatakan kesulitan dalam mengisi instrumen akreditasi; (3) sebanyak 31,11% responden menyatakan kesulitan mengisi instrumen pengumpul data dan informasi pendukung; (4) sebanyak 22,22% responden menilai kemampuan asesor dalam melakukan verifikasi dan validasi data masih kurang; (5) sebanyak 28,89% responden menilai sebagian asesor kurang mematuhi norma, tata krama, dan tata terib pelaksanaan akreditasi; (6) sebanyak 24,44% responden menilai pihak sekolah/madrasah kurang mematuhi norma, tata krama, dan tata terib pelaksanaan akreditasi; dan (7) sebanyak 60% responden mengaku rekomendasi hasil akreditasi umumnya belum ditindaklanjuti. Hasil survei ini mengindikasikan bahwa implementasi model akreditasi BAN-S/M belum sepenuhnya terlaksana secara efektif.

Permasalahan dalam penyelenggaraan akreditasi sekolah/madrasah dan implementasi model akreditasi BAN-S/M sebagaimana diuraikan di atas, berdampak pada kepercayaan *stakeholders* dan masyarakat terhadap hasil akreditasi sekolah. Masih ada pihak yang mempersepsikan bahwa pelaksanaan akreditasi kurang akuntabel dan belum sepenuhnya terpercaya. Sebagai contoh: (1)

opini yang pernah dimuat pada *Harian Lampung Pos* tanggal 27 Januari 2005 dengan judul “Menyoal Akreditasi Sekolah/Madrasah”, tulisan ini mengindikasikan bahwa objektivitas penilaian dalam proses akreditasi sekolah/madrasah masih diragukan atau kurang akuntabel; dan (2) artikel yang dimuat di *Harian Solo Pos* tanggal 23 Agustus 2007 yang berjudul “Akreditasi Benarkah Demi Peningkatan Mutu Pendidikan?”, penulis artikel mengekspresikan ketidakpercayaan terhadap proses evaluasi dalam akreditasi yang telah dilaksanakan.

Berdasarkan uraian di atas, penelitian ini bertujuan memperoleh model akreditasi SMA/MA yang efektif, efisien, dan akuntabel. Model akreditasi yang efektif, efisien, dan akuntabel diasumsikan dapat menghasilkan informasi tentang mutu sekolah/madrasah yang objektif, akurat, komprehensif, dan terpercaya yang pada akhirnya mampu mendorong dan memotivasi sekolah/madrasah untuk secara terus menerus melakukan perbaikan dan peningkatan mutu secara berkelanjutan.

Metode Penelitian

Penelitian menggunakan model penelitian dan pengembangan modifikasi gabungan model *research, development and diffusion* (model R, D & D) dari Hopkins dan Clark (Havelock, 1976) dan model *research and development* (model R & D) dari Borg & Gall (1983). Pengembangan model dilakukan dengan tahapan sebagai berikut: (1) studi permasalahan dan pengumpulan data, (2) perencanaan pengembangan model, (3) pengembangan model, (4) uji coba tahap 1, (5) uji coba tahap 2, (6) menilai keefektifan model, dan (7) penyusunan produk akhir.

Studi permasalahan dan pengumpulan data dilakukan dengan teknik wawancara dan angket. Sumber informasi adalah anggota BAP-S/M provinsi Bengkulu, asesor akreditasi SMA/MA, pengawas sekolah, kepala sekolah, guru, dan tenaga administrasi. Perencanaan pengembangan model, antara lain: (1) menyiapkan bahan referensi dan regulasi sebagai bahan rujukan untuk menyusun desain model akreditasi, (b) menen-

tukan pakar dan praktisi yang akan dilibatkan dalam pengembangan dan uji coba model, (c) memperkiraan lamanya waktu untuk kegiatan pengembangan model akreditasi, dan (d) menentukan sekolah untuk uji coba model.

Pengembangan model dilakukan mulai dari penyusunan draf model, pembahasan dan telaah draf model melalui kegiatan *focus group discussion* (FGD), dan validasi model sebelum diujicobakan. Selanjutnya dilakukan uji coba model sebanyak 2 kali. Selesai uji coba dilakukan revisi berdasarkan komentar dan saran dari praktisi. Penilaian keefektifan model dilakukan setelah selesai dilakukan uji coba. Berdasarkan hasil penilaian keefektifan model, dilakukan penyusunan produk akhir.

Pengembangan model tahap pertama adalah: (1) menentukan kriteria model akreditasi SMA/MA yang efektif, efisien, dan akuntabel; (2) menyusun aspek atau komponen penilaian akreditasi; (3) menyusun indikator kunci untuk setiap komponen standar akreditasi. FGD Tahap 1 melibatkan peserta dari unsur anggota BAP-S/M provinsi Bengkulu dan asesor akreditasi SMA/MA membahas, menelaah, dan memberikan saran perbaikan tentang kriteria model, komponen penilaian, dan rumusan indikator kunci.

Berdasarkan hasil FGD Tahap 1, dilakukan pengembangan tahap kedua: (1) membuat rancangan pengembangan instrumen akreditasi, (2) menyusun kisi-kisi pengembangan instrumen akreditasi; (3) menyusun item-item isian data dan pernyataan dalam instrumen akreditasi, (4) membuat desain prosedur akreditasi, dan (5) membuat panduan penggunaan. Pengembangan instrumen akreditasi dan prosedur akreditasi dilakukan dengan mempertimbangkan kriteria model yang efektif, efisien, dan akuntabel. FGD Tahap 2 melibatkan peserta dari unsur promotor/ kopromotor, dosen/pakar evaluasi pendidikan, dan mahasiswa S3 prodi penelitian dan evaluasi pendidikan bertujuan membahas, menelaah, dan memberi saran perbaikan terhadap desain instrumen, prosedur, dan panduan.

Instrumen akreditasi SMA/MA yang dikembangkan divalidasi oleh tiga pakar terdiri dari pakar akreditasi sekolah/madrasah, pakar manajemen pendidikan, dan pakar evaluasi pendidikan. Validasi pakar dilakukan untuk mengetahui validitas isi dan mengestimasi reliabilitas instrumen. Adapun prosedur akreditasi SMA/MA yang dikembangkan divalidasi oleh peserta FGD Tahap 2.

Data hasil FGD berupa komentar dan saran dianalisis dengan teknik deskriptif. Data hasil penilaian pakar terhadap instrumen akreditasi dianalisis dengan cara menghitung indeks koefisien V Aiken untuk mengetahui validitas isi instrumen, sedangkan estimasi reliabilitas dihitung dengan teknik *intraclass correlation coefficient (ICC)* menggunakan SPSS. Validitas prosedur akreditasi dilakukan dengan cara menghitung rerata skor berdasarkan data hasil penilaian peserta FGD.

Kriteria validitas isi instrumen disajikan pada Tabel 1. Kriteria reliabilitas instrumen disajikan pada Tabel 2. Kriteria validitas prosedur akreditasi disajikan pada Tabel 3.

Tabel 1. Kriteria Validitas Isi Instrumen

Koefisien V	Kategori
$0,75 < V \leq 1,00$	Validitas sangat baik
$0,50 \leq V \leq 0,75$	Validitas baik
$V < 0,50$	Validitas kurang baik

Tabel 2. Kriteria Reliabilitas Instrumen

ICC	Kategori
$0,75 < V \leq 1,00$	Reliabilitas sangat baik
$0,50 \leq V \leq 0,75$	Reliabilitas baik
$V < 0,50$	Reliabilitas kurang baik

Tabel 3. Kriteria Validitas Prosedur

Skor	Kategori
$X > 3,4$	Sangat baik
$2,8 < X \leq 3,4$	Baik
$2,2 < X \leq 2,8$	Cukup baik
$1,6 < X \leq 2,2$	Kurang baik
$X \leq 1,6$	Tidak baik

Uji coba dilakukan sebanyak 2 kali. Uji coba pertama dilakukan di 3 SMA dan uji coba kedua dilakukan di 6 SMA. Setiap SMA tempat uji coba diberikan satu set terdiri: panduan penggunaan, prosedur akreditasi, dan instrumen akreditasi. Subjek coba diminta mengisi instrumen akreditasi. Setelah instrumen selesai diisi, dilakukan verifikasi dan validasi isian data pada instrumen akreditasi. Selanjutnya dilakukan wawancara guna menggali informasi kelebihan dan kekurangan model yang diujicobakan. Terakhir, subjek coba diminta memberikan penilaian keefektifan model.

Subjek coba terdiri dari kepala sekolah, guru, dan tenaga administrasi dengan kriteria pernah terlibat sebagai tim penyiapan akreditasi di sekolahnya.

Data hasil wawancara berupa catatan komentar dan respon jawaban dari subjek coba dianalisis secara deskriptif. Data hasil penilaian keefektifan model dianalisis dengan cara menghitung rerata skor. Hasil perhitungan rerata skor kemudian dibandingkan dengan kriteria keefektifan seperti pada Tabel 4.

Tabel 4. Kriteria Keefektifan Model

Skor	Kategori	Kesimpulan
$X > 3,4$	Sangat baik	Dapat digunakan dan dicontoh
$2,8 < X \leq 3,4$	Baik	Dapat digunakan
$2,2 < X \leq 2,8$	Cukup baik	Dapat digunakan dengan sedikit perbaikan
$1,6 < X \leq 2,2$	Kurang baik	Dapat digunakan dengan banyak perbaikan
$X \leq 1,6$	Tidak baik	Belum dapat digunakan

Hasil Pengembangan dan Pembahasan

Hasil FGD Tahap 1

Hasil pengembangan model melalui FGD Tahap 1 diperoleh konsensus atau kesepakatan sebagai berikut: (1) kriteria model akreditasi SMA/MA yang efektif, efisien, dan akuntabel harus memenuhi kriteria tertentu; (2) urutan dan penamaan komponen

yang menjadi objek penilaian akreditasi sekolah/madrasah dapat tidak harus sama dengan urutan dan penamaan seperti urutan dan penamaan standar nasional pendidikan; dan (3) indikator kunci dikembangkan dari standar-standar akreditasi mengacu pada regulasi yang relevan dengan standar nasional pendidikan.

Kriteria model akreditasi SMA/MA yang efektif, efisien, dan akuntabel disajikan pada Tabel 5. Kriteria tersebut selanjutnya menjadi karakteristik model akreditasi SMA/MA yang dikembangkan.

Tabel 5. Kriteria Model Akreditasi SMA/MA

Kriteria	Deskripsi
Efektif	<ul style="list-style-type: none"> - Bentuk instrumen sederhana dan praktis. - Petunjuk teknis pengisian instrumen jelas dan mudah dipahami. - Instrumen dapat menjangkau data secara akurat. - Bukti-bukti pemenuhan standar akreditasi terukur. - Prosedur praktis
Efisien	<ul style="list-style-type: none"> - Pengisian data pada instrumen mudah dikerjakan. - Proses verifikasi dan validasi mudah dilakukan - Menghasilkan profil mutu sekolah secara komprehensif
Akuntabel	<ul style="list-style-type: none"> - Prosedur menjamin pihak sekolah bertindak jujur dan transparan dalam memberikan data - Prosedur menjamin asesor bertindak jujur dan objektif dalam melakukan verifikasi dan validasi data.

Berdasarkan Tabel 5, dapat dikemukakan dua aspek penting dalam model akreditasi SMA/MA, yaitu: (1) aspek instrumen, dan (2) aspek prosedur. Model akreditasi SMA/MA dengan instrumen dan prosedur akreditasi yang memenuhi kriteria efektif, efisien, dan akuntabel seperti dicantumkan pada Tabel 5 diasumsikan mampu mendorong dan memotivasi pihak sekolah untuk

melakukan upaya perbaikan dan peningkatan mutu secara berkelanjutan.

Urutan dan penamaan komponen yang menjadi objek penilaian dalam akreditasi SMA/MA disajikan pada Tabel 6.

Tabel 6. Urutan dan Penamaan Komponen Akreditasi SMA/MA

Urutan dan Nama Komponen	Acuan Regulasi
1. Standar Kompetensi Lulusan	Permendikbud No. 54 Tahun 2013 tentang Standar Kompetensi Lulusan
2. Standar Penilaian	Permendikbud No. 66 Tahun 2013 tentang Standar Penilaian
3. Standar Proses Pembelajaran	Permendikbud No. 65 Tahun 2013 tentang Standar Proses
4. Standar Isi Kurikulum	Permendikbud No. 64 Tahun 2013 tentang Standar Isi
5. Standar Pendidik dan Tenaga Kependidikan	Permendiknas No. 13 Tahun 2006 tentang Standar Kepala Sekolah; Permendiknas No. 16 Tahun 2006 tentang Standar Guru; Permendiknas No. 27 Tahun 2008 tentang Standar Konselor; Permendiknas No. 24 Tahun 2008 tentang Standar Tenaga Administrasi; Permendiknas No. 25 Tahun 2008 tenaga Perpustakaan; dan Permendiknas No. 26 Tahun 2008 tentang Standar Tenaga Laboratorium.
6. Standar Sarana dan Prasarana	Permendiknas No. 24 Tahun 2007 tentang Standar Sarana dan Prasarana
7. Standar Pengelolaan	Permendiknas No. 19 Tahun 2007 tentang Standar Pengelolaan
8. Standar Pembiayaan	Permendiknas No. 69 Tahun 2009 tentang Standar Biaya

Tabel 6 menunjukkan urutan dan penamaan komponen akreditasi sedikit berbeda bila dibanding dengan urutan dan penamaan model yang sudah ada, yakni model akreditasi BAN-S/M. Urutan dan penamaan dalam model akreditasi SMA/MA yang

dikembangkan diasumsikan bersifat lebih generik dan tetap mencakup aspek mutu sekolah yang bersifat komprehensif. Selain itu, acuan regulasi yang digunakan pengembangan instrumen akreditasi disesuaikan dengan regulasi yang terbaru yakni regulasi tentang: standar kompetensi lulusan, standar penilaian, standar isi, dan standar proses.

Regulasi sebagaimana tercantum pada 6 digunakan sebagai acuan dalam pengembangan indikator kunci untuk setiap komponen standar akreditasi.

Contoh indikator kunci untuk standar akreditasi dalam model akreditasi SMA/MA yang dikembangkan disajikan pada Tabel 7.

Tabel 7. Indikator Kunci untuk Standar Akreditasi 1. Kompetensi Lulusan

No. Item	Indikator Kunci
1.1	Lulusan memiliki perilaku yang mencerminkan sikap orang beriman, berakhlak mulia, berilmu, percaya diri, dan bertanggung jawab.
1.2	Lulusan memiliki pengetahuan faktual, konseptual, prosedural, dan metakognitif dalam ilmu pengetahuan, teknologi, seni, dan budaya.
1.3	Lulusan memiliki kemampuan pikir dan tindak yang efektif dan kreatif dalam ranah abstrak dan konkret.
1.4	Lulusan melanjutkan perguruan tinggi.

Indikator kunci dari setiap komponen standar akreditasi digunakan sebagai acuan dalam pengembangan instrumen akreditasi. Prinsip pengembangan instrumen diselaraskan dengan kriteria model akreditasi SMA/MA yang efektif dan efisien.

Hasil FGD Tahap 2

Hasil pengembangan model melalui FGD Tahap 2 diperoleh desain model akreditasi SMA/MA terdiri dari: (1) instrumen akreditasi, (2) prosedur akreditasi, dan (3) panduan penggunaan. Instrumen akreditasi dalam model akreditasi SMA/MA yang dikembangkan divalidasi oleh tiga pakar yaitu: pakar akreditasi sekolah/madrasah, pakar manajemen pendidikan, dan pakar evaluasi pendidikan.

Hasil analisis validitas isi instrumen dengan cara menghitung indeks koefisien V Aiken. Koefisien validitas aiken berkisar antara 0,76 sampai dengan 0,92, dengan rata-rata mencapai 0.81 (lihat tabel 8). Hal ini menunjukkan bahwa instrumen memiliki validitas isi termasuk kategori sangat baik.

Tabel 8. Hasil Analisis Koefisien Aiken'V

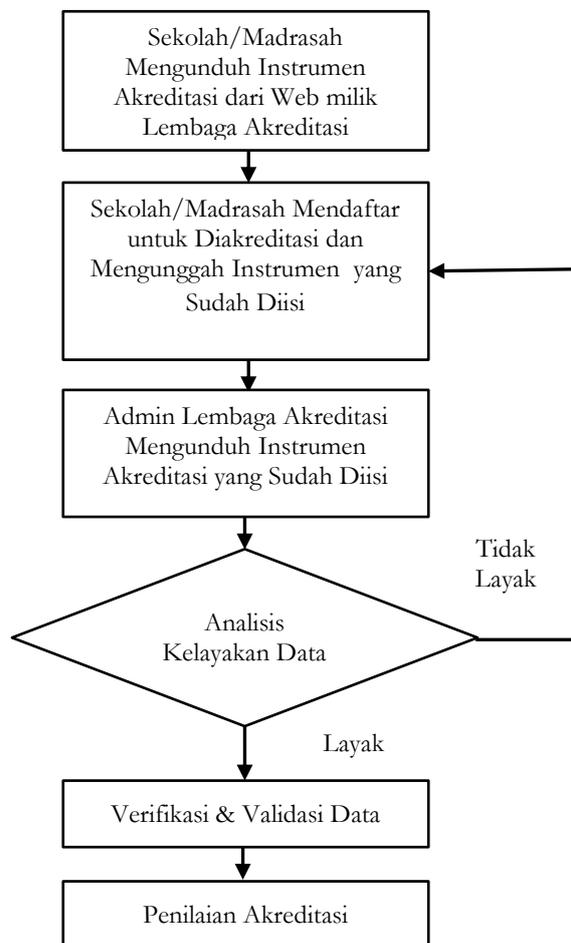
Komponen Penilaian Akreditasi	Rerata Indeks Koefisien V
Standar Akreditasi 1. Kompetensi Lulusan	0,81
Standar Akreditasi 2. Penilaian	0,81
Standar Akreditasi 3. Proses Pembelajaran	0,79
Standar Akreditasi 4. Isi Kurikulum	0,80
Standar Akreditasi 5. Pendidik dan Tenaga Kependidikan	0,82
Standar Akreditasi 6. Sarana dan Prasarana	0,92
Standar Akreditasi 7. Pengelolaan	0,76
Standar Akreditasi 8. Pmbiayaan	0,81
Rerata Indeks Koefisien Aiken'V Keseluruhan	0,81

Hasil analisis estimasi reliabilitas instrumen dengan teknik *interclass correlation coefficient* (ICC) menggunakan SPSS. Hasil analisis ICC dari ketiga rater (penilai) sebesar 0,75. Hal ini menunjukkan bahwa instrumen mempunyai reliabilitas termasuk kategori baik (lihat Tabel 9).

Tabel 9. Hasil Analisis Interclass Correlation Coefficient (ICC)

	ICC	95% Confidence Interval		F Test with True Value 0			
		Lower Bound	Upper Bound	Value	df1	df2	Sig
Single Measures	.506 ^a	.348	.651	4.070	54	108	.000
Average Measures	.754 ^c	.616	.849	4.070	54	108	.000

Hasil pengembangan prosedur akreditasi diperoleh prosedur dengan langkah-langkah seperti terlihat pada Gambar 1.



Gambar 1. Prosedur Akreditasi SMA/MA yang Dikembangkan

Pihak sekolah mengunduh instrumen akreditasi yang tersedia di web milik lembaga akreditasi. Pihak sekolah mengisi instrumen akreditasi dalam bentuk aplikasi *microsoft excel*. Isian data akreditasi ini *hyper-link* dengan: (1) instrumen pernyataan capaian standar akreditasi dan (2) hasil penilaian akreditasi. Selanjutnya pihak sekolah mendaftar untuk diakreditasi dan mengunggah instrumen akreditasi yang sudah diisi disertai file dalam bentuk dokumen, foto, atau video.

Langkah berikutnya, admin yang bekerja pada lembaga akreditasi mengunduh instrumen akreditasi dari sekolah yang akan diakreditasi. Data yang diperoleh dianalisis kelayakannya. Analisis dikerjakan oleh asesor.

Apabila hasil analisis diperoleh keputusan layak untuk diakreditasi, maka lembaga akreditasi menugaskan asesor untuk melakukan visitasi guna memverifikasi dan memvalidasi data yang telah diberikan oleh pihak sekolah/madrasah. Langkah terakhir adalah penilaian akreditasi berdasarkan data, bukti, dan catatan hasil visitasi yang dikumpulkan oleh asesor.

Prosedur akreditasi seperti dipaparkan di atas mendapat penilaian baik dari peserta FGD. Tabel 10 disajikan hasil penilaian (validasi) prosedur akreditasi. Berdasarkan Tabel 10 diperoleh informasi rerata skor keseluruhan sebesar 3,16 termasuk kategori baik.

Tabel 10. Hasil Penilaian Prosedur Akreditasi

Aspek yang Dinilai	Rerata Skor
Prosedur praktis	3,35
Prosedur menjamin pihak sekolah bertindak jujur dan transparan. .	3,32
Prosedur menjamin asesor bertindak jujur, adil, dan objektif.	2,91
Proses verifikasi dan validasi mudah dikerjakan	2,96
Fisibilitas (dapat diterapkan)	3,26
Rerata Skor Keseluruhan	3,16

Tabel 11. Hasil Penilaian Keefektifan Model

Aspek yang Dinilai	Rerata Skor
Bentuk instrumen sederhana	3,39
Keterbacaan petunjuk teknis pengisian instrumen	3,30
Keakuratan data	3,43
Kesesuaian data dengan bukti pendukung	3,39
Prosedur praktis	3,52
Kemudahan pengisian data	3,22
Kemudahan melakukan verifikasi dan validasi data	3,17
Komponen penilaian komprehensif	3,39
Kepercayaan dapat diterapkan secara jujur dan objektif	3,22
Fisibilitas (kemungkinan dapat digunakan)	3,52
Kejelasan panduan penggunaan	3,48
Rerata Skor Keseluruhan	3,40

Hasil uji coba, diperoleh keefektifan model berdasarkan penilaian dari kepala sekolah, guru, dan tenaga administrasi di sekolah tempat uji coba yang disajikan pada Tabel 11. Berdasarkan Tabel 11 diperoleh informasi rerata skor keefektifan model sebesar 3,40 termasuk kategori baik. Dengan demikian dapat disimpulkan bahwa model akreditasi SMA/MA yang dikembangkan dapat digunakan.

Simpulan

Berdasarkan hasil penelitian dan pembahasan yang telah disajikan, maka dapat disampaikan simpulan sebagai berikut. Produk yang dikembangkan memiliki karakteristik sebagai model akreditasi SMA/MA yang efektif, efisien, dan akuntabel yaitu: (1) bentuk instrumen sederhana/praktis, (2) petunjuk teknis pengisian instrumen jelas dan mudah dipahami, (3) instrumen dapat menjaring data secara akurat, (4) bukti pendukung data terukur, (5) prosedur praktis, (6) pengisian data mudah dikerjakan, (7) verifikasi dan validasi data mudah dilakukan, (8) menghasilkan profil mutu sekolah secara komprehensif, (9) prosedur menjamin pihak sekolah bertindak jujur dan transparan; dan (10) prosedur menjamin asesor bertindak jujur dan objektif dalam melakukan verifikasi dan validasi data. Model akreditasi SMA/MA yang dikembangkan ini diberi nama Model Akreditasi SMART.

Instrumen akreditasi dalam Model Akreditasi SMART memiliki validitas isi berdasarkan indeks koefisien Aiken's V sebesar 0,81 termasuk kategori sangat baik dan reliabilitas berdasarkan *Interclass Correlation Coefficient* (ICC) sebesar 0,75 termasuk kategori baik.

Prosedur akreditasi dalam Model Akreditasi SMART terdiri dari enam langkah yaitu: (1) sekolah/madrasah mengunduh instrumen akreditasi dari web milik lembaga akreditasi (*online*), (2) sekolah/madrasah mendaftar untuk diakreditasi dan mengunggah instrumen yang sudah diisi (*online*), (3) admin lembaga akreditasi mengunduh instrumen akreditasi yang sudah diisi (*online*), (4) analisis kelayakan data, (5) veri-

fikasi dan validasi data melalui visitasi, dan (6) penilaian akreditasi.

Prosedur akreditasi dalam Model Akreditasi SMART mendapat penilaian dari peserta FGD dengan rerata skor sebesar 3,16 termasuk kategori baik. Model Akreditasi SMART mendapat penilaian dari kepala sekolah, guru, dan tenaga administrasi dengan rerata skor sebesar 3,12 termasuk kategori baik dan model dapat digunakan.

Berdasarkan simpulan penelitian tersebut, maka penelitian ini dapat memberikan saran sebagai berikut. Pihak SMA/MA dapat memanfaatkan instrumen dalam Model Akreditasi SMART sebagai alat dalam melakukan evaluasi diri sekolah (EDS). Dinas Pendidikan, Kantor Kemenag, dan Pengurus Yayasan penyelenggara pendidikan dapat memanfaatkan Model Akreditasi SMART sebagai alat pengendalian dan evaluasi mutu SMA/MA yang menjadi kewenangannya. LPMP dapat memanfaatkan Model Akreditasi SMART sebagai alat dalam melakukan pemetaan mutu SMA/MA dalam kerangka penjaminan mutu. BAN-S/M dapat mengadopsi, mengadaptasi, dan memanfaatkan Model Akreditasi SMART sebagai bahan referensi dalam pengembangan Model Akreditasi BAN-S/M. Pakar evaluasi pendidikan dan peneliti lainnya dapat menggunakan Model Akreditasi SMART sebagai bahan rujukan dan referensi untuk keperluan penelitian sejenis dan/atau mengembangkan hasil penelitian pengembangan ini lebih lanjut.

Daftar Pustaka

- Artcraft. (2005). *Program evaluation*. Sydney: Artcraft Research.
- Aswandi. (2009). Akreditasi sekolah 2009. Retrieved December 25, 2010, from <http://www.pontianak/post.com/?mb=berita.detail&id=27077>
- Bennett, W. . (2000). *Independen school accreditation: A guide to school choice*. Retrieved December 22, 2008, from Error! Hyperlink reference not valid.

- Borg, W. R., & Gall, M. D. (1983). *Educational research: An introduction* (4th ed.). New York: Longman Publishing.
- COA. (2008). Accreditation handbook: definition of stipulations 2007-08. Retrieved July 19, 2008, from https://www.ctc.ca.gov/docs/default-source/educator-prep/coa-agendas/2008-05/2008-05-item-16.pdf?sfvrsn=ba0760a2_0
- Depdiknas. (2002). *Manajemen peningkatan mutu berbasis sekolah: konsep dasar*. Jakarta: Depdiknas.
- ECA. (2005). Accreditation in the European higher education area. Retrieved July 19, 2008, from http://www.aic.lv/bologna/Bologna/contrib/Statem_oth/ECA_on_Bergen.pdf
- Griffin, P., & Nix, P. (1991). *Educational assessment and reporting*. Sydney: Harcourt Brace Javanovic Publisher.
- Havelock, R. G. (1976). *Planning for innovation through dissemination and utilization of knowledge*. Michigan: Institute for Social Research The University of Michigan.
- Huitt, W. (2004). *Assessment, measurement, and evaluation. Educational Psychology Interactive*. Valdosta, GA: Valdosta State University. Retrieved from <http://www.edpsycinteractive.org/topics/measeval/measeval.html>
- Kartowagiran, B. (2006). *Prinsip-prinsip dasar monitoring dan implementasinya. Bahan Penyelenggaraan Calon Tim Pelatih Monitoring dan Evaluasi Provinsi DKI Jakarta*. Yogyakarta: Universitas Negeri Yogyakarta.
- Lindsay, D. H., & Campbell, A. (2002). An examination of AACSB accreditation status an accounting program quality indikator. Retrieved October 22, 2008, from <http://www.sbaer.uca.edu/research/TemporarilyDisabled---wdsi/2002/pdffiles/papers/012.pdf.htm>
- Mardapi, D. (2003). *Konstruksi tes dan analisis butir* (Makalah tidak dipublikasikan). Yogyakarta.
- Mardapi, D. (2008). *Teknik penyusunan instrumen tes dan nontes*. Yogyakarta: Mitra Cendikia Press.
- Menteri Pendidikan Nasional. Peraturan pemerintah republik indonesia nomor 19 tahun 2005 tentang standar nasional pendidikan (2005).
- Menteri Pendidikan Nasional. (2007). *Pedoman pelaksanaan akreditasi sekolah/madrasah*. Jakarta: Badan Akreditasi Sekolah/Madrasah.
- Menteri Pendidikan Nasional. Peraturan Mendiknas Nomor 63 Tahun 2009 tentang Sistem Penjaminan Mutu Pendidikan (2009).
- Pagliarulo, M. A. (1996). Accreditation: Its nature, process, and effective implementation. *Physical Therapy Education Journal*, 66(7), 1114–1118.
- Patil, A. S., & Pudlowski, Z. J. (2005). Important issues of the accreditation and quality assurance and a strategy in the development of an accreditation framework for engineering courses. *Global Journal of Engineering Education*, 9(1), 49–58.
- Presiden Republik Indonesia. Undang-Undang Republik Indonesia nomor 20 tahun 2003 tentang sistem pendidikan nasional, Pub. L. No. 20 (2003). Indonesia.
- Stark, J. S., & Thomas, A. (1994). *Assessment and program evaluation*. Needham Heights: Simon and Schuster Custom Publishing.
- Stufflebeam, D. L., & Shinkfield, A. J. (2007). *Evaluation, theory, models, and applications*. San Francisco: John Wiley & Sons, Inc.
- Sumarno. (2000). *Pengembangan model akreditasi sekolah negeri tingkat SLTP dan SMU*. Yogyakarta: Universitas Negeri Yogyakarta.

- Sutadji, E. (2009). *Model evaluasi mutu sekolah: pengembangan instrumen untuk menetapkan mutu*. Disertasi tidak dipublikasikan. Program Pascasarjana Universitas Negeri Yogyakarta.
- Thorndike, R. M., Cunningham, G. K., Thorndike, R. L., & Hagen, E. P. (1991). *Measurement and evaluation in psychology and education* (5th ed.). New York: Macmillan Publishing Co, Inc.
- Worthen, B. R., & Sander, J. R. (1989). *Educational evaluation: theory and practice*. Worthington, Ohio: Charles A. Jones Publishing Company.