

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/359920276>

Model–Model Evaluasi Pendidikan Dasar (Scriven Model, Tyler Model, dan Goal Free Evaluation)

Article in Jurnal Pendidikan Riset dan Konseptual · January 2022

DOI: 10.28926/riset_konseptual

CITATIONS

0

READS

1,855

3 authors, including:



Mami Hajaroh

Universitas Negeri Yogyakarta

37 PUBLICATIONS 68 CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:



Pendidikan Berbasis Nilai di Taman Kanak-Kanak [View project](#)



Difusi Kebijakan Gender Maisntreaming dalam Organisasi Fatayat NU [View project](#)

Model-Model Evaluasi Pendidikan Dasar (Scriven Model, Tyler Model, dan Goal Free Evaluation)

Helda Kusuma Wardani⁽¹⁾, Fajarsih Darusuprpti⁽²⁾, Mami Hajaroh⁽³⁾

^{1,2,3} Universitas Negeri Yogyakarta, Indonesia

Email: ¹heldakusuma.2021@student.uny.ac.id,

²fajarsihdarusuprpti.2021@student.uny.ac.id, ³mami_hajaroh@uny.ac.id

Abstrak: Evaluasi program menjadi penting untuk dipahami dan dilakukan pendidik, karena setiap pendidik harus mampu menjadi evaluator. Model evaluasi program banyak ragam dan fungsinya tergantung dari jenis program yang dievaluasi. Dalam pendidikan sedikitnya ada tiga jenis program, yakni program pemrosesan, pelayanan, dan umum. Evaluasi program pemrosesan cocok dievaluasi dengan *Formative-Summative Evaluation Model* atau Scriven Model, GOE atau Tyler Model, dan GFE Model. Dua model evaluasi program yaitu Scriven model dan Tyler model digunakan di Indonesia. Kedua model ini seringkali digunakan oleh guru saat merancang, melaksanakan, dan evaluasi program pembelajarannya.

Tersedia Online di

http://journal.unublitar.ac.id/pendidikan/index.php/Riset_Konseptual

Sejarah Artikel

Diterima pada : 21-11-2021

Disetujui pada : 28-01-2022

Dipublikasikan pada : 28-01-2022

Kata Kunci:

Evaluasi Program, Scriven Model, Tyler Model

DOI:

http://doi.org/10.28926/riset_konseptual.v6i1.446

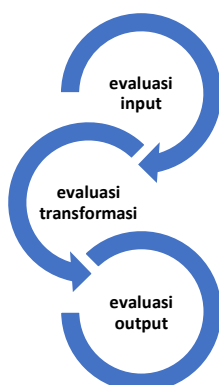
PENDAHULUAN

Proses siklus terjadi dalam pengembangan suatu program pendidikan atau program pembelajaran, dimulai dari perancangan, pengembangan, dan evaluasi kembali pada aktivitas perancangan. Setiap siklus diharapkan memperoleh data yang dapat digunakan untuk perbaikan program supaya lebih efisien atau untuk menentukan keefektifan program oleh pembuat keputusan. Model evaluasi yang digunakan tergantung pada jenis program yang dievaluasi. Ada tiga macam program seperti dikemukakan oleh Suharsimi & Cepi (2010) meliputi program (1) pemrosesan, (2) layanan, dan (3) umum (Arikunto & Jabar, 2010, pp. 49-52)

Program pemrosesan merupakan rancangan yang memuat aktivitas untuk mengubah input menjadi output yang diinginkan, sehingga membentuk suatu proses sistematis dan tersistem. Program pemrosesan juga disebutkan sebagai program dengan aktivitas inti memproses masukan mentah menjadi hasil atau produk atau keluaran (Arikunto & Jabar, 2010, p. 49) Contoh program pemrosesan yang ada di pendidikan SD diantaranya program pembelajaran, program Penguatan Pendidikan Karakter, program GLS, program kepramukaan.

Obyek evaluasi program pemrosesan tentunya mencakup input, transformasi atau proses, dan output. Input sebagai obyek evaluasi dapat berupa kondisi awal siswa meliputi perilaku masukan (entry behavior), motivasi, ataupun inteligensi. Sedangkan obyek evaluasi pada proses transformasi meliputi kurikulum, isi, teknologi, pedagogi ataupun guru dan tenaga kependidikan. Obyek output dapat berupa hasil belajar, perubahan karakter, kemampuan literasi bahasa, literasi informasi, ataupun literasi numerasi (Panjaitan, 2009, p. 35) Bagaimanakah posisi obyek evaluasi program ini? Akankah evaluasi program pemrosesan dilakukan dengan mengabaikan obyek evaluasi program?

Evaluasi program pemrosesan seperti yang diutarakan oleh Suharsimi dan Cepi (2010) dan obyek evaluasi program pemrosesan dapat digambarkan secara diagramatik seperti berikut.



Gambar 1: Diagram Proses Evaluasi Program Pemrosesan

Gambar 1 yang merupakan diagram proses evaluasi program pemrosesan memunculkan pertanyaan “model evaluasi program apakah yang sesuai untuk evaluasi program pemrosesan?”. Untuk mengevaluasi program pemrosesan ada tujuh model evaluasi program yang dapat digunakan, meliputi: (1) *Goal Oriented Evaluation (GOE) Model*, (2) *Goal Free Evaluation (GFE) Model* (3) *Formative-Summative Evaluation Model*, (4) Model Deskripsi Pertimbangan, (5) Model Evaluasi CSE_UCLA, (6) Model Evaluasi CIPP, dan (7) Model Evaluasi Kesenjangan (Arikunto & Jabar, 2010, 52-55) Apakah tujuh model evaluasi program ini mampu mengevaluasi program pemrosesan? Dari tujuh model evaluasi program pemrosesan, tidak semuanya akan dibahas dalam kajian pustaka ini. Ada 3 (tiga) yang pertama dari model evaluasi program pemrosesan yang akan dibahas. Yaitu Scriven model, Tyler model dan Goal Free Evaluation.

PEMBAHASAN

A. Model Evaluasi Program

Pada awal pengembangan lapangan, Scriven (1967) mendefinisikan evaluasi sebagai menguji nilai atau manfaat dari program (Fitzpatrick, Sanders, & Worthen, 2011, p. 7). Banyak pengertian terbaru mencakup definisi asli dari istilah ini. Banyak yang setuju bahwa evaluasi menguji nilai atau manfaat dari objek evaluasi. Salah satu peran penting evaluator yaitu membantu para pemangku kepentingan mengartikulasikan kriteria dan untuk merangsang dialog tentang kriteria ini.

Definisi evaluasi menekankan menggunakan kriteria untuk menguji manfaat atau nilai produk, yang dapat dijabarkan sebagai metode penyelidikan dan keputusan, termasuk: (1) ditetapkannya kriteria dan standar untuk menilai kualitas dan memutuskan apakah standar tersebut harus relatif atau absolut, (2) mengumpulkan informasi yang relevan, dan (3) menerapkan standar untuk menentukan nilai, kualitas, utilitas, efektivitas, atau signifikansi. Ini mengarah pada rekomendasi yang dimaksudkan untuk mengoptimalkan objek evaluasi dalam kaitannya dengan tujuan yang dimaksudkan atau untuk mendukung para pemangku kepentingan menentukan program ditolak atau dilanjutkan (Fitzpatrick, Sanders, & Worthen, 2011, p. 7). Program merupakan salah satu yang menjadi obyek evaluasi. Selaras dengan Ornstein (2018) yang menyatakan evaluasi adalah proses di mana orang mengumpulkan data untuk membuat keputusan. Terlepas dari keumuman itu, bagaimanapun, definisi evaluasi bervariasi. Ada yang mengemukakan evaluasi sebagai cara-cara formal menentukan kualitas atau nilai program pendidikan atau kurikulum. Cakupan evaluasi berupa aktivitas penilaian yang meliputi: (1) menentukan standar yang normatif atau patokan, (2) pengumpul dan data an informasi yang dibutuhkan, dan (3) menentukan kualitas atau nilai (Ornstein & Hunkins, 2018, p. 294)

Program secara kebahasaan dapat diartikan sebagai suatu rancangan usaha yang akan dijalankan, ini maknanya program ada sejak sebelum, selama, dan sesudah aktivitas yang dirancang. Komite Bersama Standar Evaluasi Pendidikan (1994) mendefinisikan program hanya sebagai "kegiatan yang disediakan secara

berkelanjutan". Dalam edisi baru Standar mereka (2010) Komite Bersama mencatat bahwa sebuah program jauh lebih dari satu seperangkat kegiatan. Mereka menulis definisi sepenuhnya, sebuah program adalah Seperangkat kegiatan sistematis yang direncanakan menggunakan sumber daya terkelola untuk mencapai tujuan tertentu terkait dengan kebutuhan spesifik individu atau kelompok manusia yang spesifik, diidentifikasi, berpartisipasi dalam konteks tertentu menghasilkan output, hasil, dan dampak yang dapat didokumentasikan mengikuti sistem keyakinan yang diasumsikan (eksplisit atau implisit) (diagnostik, kausal, intervensi, dan teori implementasi tentang cara kerja program) . Dengan spesifik, biaya dan manfaat yang dapat diinvestasikan (Fitzpatrick, Sanders, & Worthen, 2011, p. 8)

Evaluasi program merupakan suatu proses menentukan nilai dan manfaat dari rancangan usaha yang akan dijalankan. Evaluasi program dapat juga diartikan sebagai proses pengumpulan data sebelum, selama, dan sesudah usaha yang dirancang diimplementasikan untuk menentukan nilai dan manfaatnya. Hasil analisis data yang dikumpulkan dapat digunakan untuk memperbaiki ataupun untuk menentukan keefektifannya.

Model dapat diartikan sebagai representasi procedural dari aktivitas yang dilakukan oleh manusia, dapat diwujudkan dalam bagan alur atau deskripsi proseduralnya. Dengan demikian, model evaluasi program adalah representasi procedural proses pengumpulan data sebelum, selama, dan sesudah usaha yang dirancang diimplementasikan untuk menentukan nilai dan manfaatnya. Model evaluasi program ditentukan oleh pengembangnya masing-masing, seperti Scriven Model, Tyler Model, dan GFE Model yang akan dibahas satu persatu berikut ini.

B. Scriven Model

1. Konsepsi Scriven Model

Michael Scriven lengkapnya Michael John Scriven seorang filsuf akademis dan akademisi yang terkenal karena kontribusinya dalam teori dan praktik evaluasi. Salah satu yang sering disebut sebagai Scriven Model adalah model evaluasi formatif dan sumatif.

Evaluasi formatif merupakan evaluasi yang dilakukan selama perancangan dan pengembangan program, dengan maksud memberikan umpan balik untuk meningkatkan evaluasi (obyek yang dievaluasi), dapat juga fokus pada rencana program atau desain (Mertens & Wilson, 2019, p. 545). Pada evaluasi formatif ini pengumpulan data dan informasi diarahkan pada kekurangan atau kesalahan yang ada dalam program yang dievaluasi. Data dan informasi yang dikumpulkan akan digunakan untuk perbaikan program atau produk sebagaimana ataupun secara keseluruhan. Evaluator hanya bertindak untuk memperoleh keputusan hasil evaluasi, sedangkan tindak lanjut hasil evaluasi dikembalikan kepada *stakeholder*.

Evaluasi formatif dilaksanakan sebelum program dilaksanakan maupun selama program dilaksanakan, dengan perbedaan pada evaluator dan data atau informasi yang dikumpulkan untuk dianalisa dan dibuat rekomendasinya kepada *stakeholder*. Sebelum program dilaksanakan, evaluator yang dilibatkan adalah konsultan ahli dan pengguna atau sasaran program. Evaluator konsultan ahli disesuaikan dengan jenis program yang dievaluasi, misal untuk program pembelajaran yang melibatkan ahli matapelajaran (*Subject Matter Expert*), special pembelajaran (*Learning Specialist*), spesialis pembelajar (*Learner Specialist*). Sedangkan pengguna atau sasaran program yang terlibat adalah pendidik profesional (guru) dan peserta didik dan/atau orangtua (bila dilibatkan dalam program).

Data minimal yang seharusnya dikumpulkan untuk dianalisis meliputi data dan informasi berikut.

- Reaksi ahli (pakar) bidang sesuai program, yang bertanggung jawab untuk memverifikasi bahwa konten dan konstruk.

- Reaksi seorang manajer atau supervisor yang telah mengamati pelajar menggunakan keterampilan dalam konteks kinerja.
- Program pembelajaran membutuhkan data uji lapangan meliputi data kuantitatif hasil tes, dan data kualitatif hasil pengamatan, wawancara, ataupun kuesioner. Juga data tentang waktu yang dibutuhkan untuk menyelesaikan program.

Keseluruhan data dan informasi kualitatif maupun kuantitatif ini akan membantu perbaikan program sebelum diimplementasikan.

Peran dari masing-masing pihak eksternal dalam penilaian formatif ini dibutuhkan untuk perbaikan program, sehingga kejelasan, kecermatan, dan ketepatan data serta informasi sangat dibutuhkan. Data dan informasi dari para pakar tidak selalu digunakan oleh perancang program, sebelum program diimplementasikan kepada pengguna (guru dan peserta didik) (Dick, Carey, & Carey, 2015, pp. 287-288). Bisa saja data dan informasi tentang kesalahan dan kekurangan program digunakan sebelum ujicoba lapangan secara terbatas.

Evaluasi sumatif hanya dilakukan setelah program dilaksanakan untuk mengumpulkan data dan informasi yang dapat dianalisis dalam menentukan nilai keefektifan & efisiensi kemanfaatannya. Proses mengumpulkan data dan informasi untuk membuat keputusan tentang program benar-benar bekerja sebagaimana dimaksud dalam konteks kinerja. Lebih lanjut, ini digunakan untuk menentukan apakah kemajuan sedang dibuat dalam memperbaiki masalah kinerja yang mendorong upaya desain dan pengembangan instruksional. Tujuan utama evaluasi sumatif untuk menentukan apakah program memenuhi harapan (Dick, Carey, & Carey, 2015, p. 343).

Data dan informasi yang dikumpulkan dalam evaluasi sumatif program minimal meliputi berikut ini.

- Data dan informasi kondisi sasaran atau pengguna sebelum program diselenggarakan.
- Data dan informasi kondisi sasaran atau pengguna sesudah program diselenggarakan.
- Data dan informasi kondisi selama proses atau transformasi program diselenggarakan.
- Data dan informasi keefektifan ataupun efisiensi penyelenggaraan program.

Data dan informasi inilah yang kemudian dianalisis dan diinterpretasikan oleh para ahli untuk menentukan nilai dan manfaat program yang diselenggarakan.

Ada dua aktivitas evaluasi sumatif, yaitu: (1) penilaian ahli, dan (2) analisis dampak. Penilaian ahli dapat dilaksanakan melalui berbagai cara berikut.

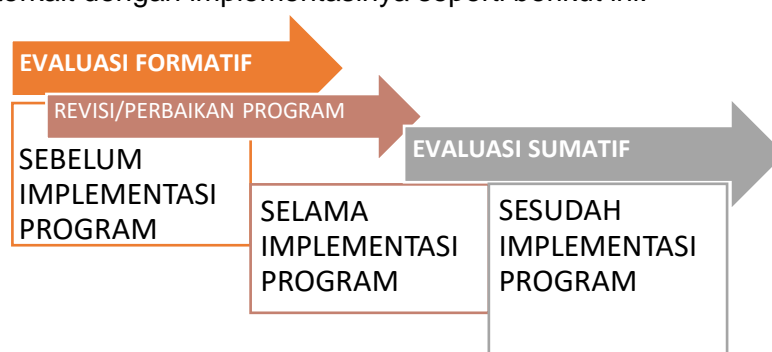
- (1) Analisis kongruensi, yakni membandingkan tujuan dari lembaga dengan tujuan yang dirumuskan dalam program.
- (2) Analisis konten, yakni evaluasi akurasi, kekinian, dan keakuratan konten yang ada dalam program.
- (3) Analisis desain, yakni evaluasi terhadap kecukupan strategi implementasi program.
- (4) Analisis kelayakan transfer, yakni evaluasi terhadap kelayakan program mentransfer konten kepada sasaran.

Evaluasi sumatif yang dirancang dengan baik akan memperoleh fungsi untuk: (1) menyediakan bukti sebuah hubungan kausalitas, (2) menilai dampak jangka panjang, dan (3) menyediakan data mengenai dampak program. Secara keseluruhan fungsi evaluasi formatif-sumatif untuk memperbaiki program sebelum dan selama implementasi program, dan untuk menentukan angka kemajuan atau capaian hasil sasaran program. Tujuan model evaluasi formatif - sumatif atau Scriven Model terhadap program berupa kurikulum, meliputi: (1) Menentukan hasil dari suatu program, (2) Membantu dalam memutuskan apakah akan menerima atau menolak program, (3) Pastikan perlunya revisi konten kursus, (4) Membantu dalam pengembangan materi kurikulum di masa depan untuk perbaikan berkelanjutan, dan

(5) Meningkatkan metode pengajaran dan teknik instruksional (Glatthorn, Boschee, Whitehead, & Boschee, 2019, p. 500).

2. Prosedur Scriven Model

Perbedaan peran evaluasi formatif dan evaluasi sumatif dilakukan Scriven pertama kali tahun 1967. Sejak itu, istilah-istilah tersebut telah diterima di lapangan. Dalam praktiknya, perbedaan antara kedua jenis evaluasi ini mungkin agak kabur, tetapi istilah tersebut berfungsi sebagai fungsi penting dalam menyoroti jenis keputusan atau pilihan yang dapat dilayani oleh evaluasi. Istilah-istilah tersebut, pada kenyataannya, kontras dua jenis tindakan yang mungkin diambil oleh para pemangku kepentingan sebagai hasil dari evaluasi (Fitzpatrick, Sanders, & Worthen, 2011, p. 20). Kekontrasan dari keduanya, evaluasi formatif dan evaluasi sumatif, seringkali digunakan sebagai dasar untuk memisahkan dan membedakan kedua aktivitas tersebut walaupun menjadi satu kesatuan Scriven Model. Namun demikian, dapat digambarkan dalam bagan alur terjadinya evaluasi formatif dan evaluasi sumatif terkait dengan implementasinya seperti berikut ini.



Gambar 2: Bagan alur Prosedural Tahapan Scriven Model

Tahapan pertama terdiri dua aktivitas secara sekuensial-simultan yang terdiri dari evaluasi formatif program dan revisi/perbaikan program dilakukan sebelum dan selama implementasi program. Evaluasi sumatif dilaksanakan secara simultan-sekuensial selama dan sesudah implementasi program. Berdasarkan bagan alur procedural tersebut, dapat dijabarkan kegiatan Scriven Model seperti berikut.

Tahap pertama: Evaluasi Formatif dan Revisi Program

- (1) Memberikan program kepada pakar atau konsultan ahli untuk dievaluasi kekurangan dan kesalahan yang pada program sebelum implementasi program. Hasil dari evaluasi ahli ini dapat digunakan untuk revisi ataupun tidak, tergantung dari perancang program.
- (2) Evaluasi formatif dengan pengguna atau sasaran program serempak dengan implementasi program, hasil evaluasi formatif pengguna/sasaran program bersama evaluasi formatif ahli digunakan untuk revisi perbaikan program.
- (3) Evaluasi formatif ahli diikuti evaluasi formatif pengguna sasaran (dilakukan secara sekuensial), dan serentak dengan kedua evaluasi formatif dilakukan revisi perbaikan program (dilakukan secara simultan). Sehingga prosedur tahap pertama ini dapat disebut sebagai tahap sekuensial-simultan.

Tahap kedua: Evaluasi Sumatif

- (1) Saat revisi perbaikan program masih dilaksanakan dan akan berakhir, dimulailah kegiatan evaluasi sumatif selama implementasi program (dilakukan secara simultan).
- (2) Setelah implementasi program selesai, evaluasi sumatif tetap dilanjutkan setelah implementasi program (dilakukan secara sekuensial).
Oleh karena ada kegiatan evaluasi sumatif dilakukan simultan dengan revisi program kemudian dilanjutkan dengan evaluasi sumatif, maka tahap kedua ini merupakan aktivitas simultan-sekuensial.

3. Implementasi Scriven Model

Scriven model merupakan model evaluasi program yang secara yuridis formal digunakan dalam standar penilaian dan tertuang dalam bagian kelima PPRI Nomor 57 Tahun 2021 tentang Standar Nasional Pendidikan. Pada pasal 16 ayat 5 dinyatakan bahwa Penilaian hasil belajar Peserta Didik sebagaimana dimaksud pada ayat (4) berbentuk: a. penilaian formatif; dan b. penilaian sumatif. (Presiden, 2021) Dalam pasal 17 dan 18 dijabarkan yang dimaksud dengan penilaian formatif dan penilaian sumatif hasil belajar. Sedangkan pasal 16 ayat (2) Mekanisme sebagaimana dimaksud pada ayat (1) merupakan prosedur dalam melakukan penilaian yang meliputi:

- a. perumusan tujuan penilaian;
- b. pemilihan dan/atau pengembangan instrument penilaian;
- c. pelaksanaan penilaian;
- d. pengolahan hasil penilaian; dan
- e. pelaporan hasil penilaian.

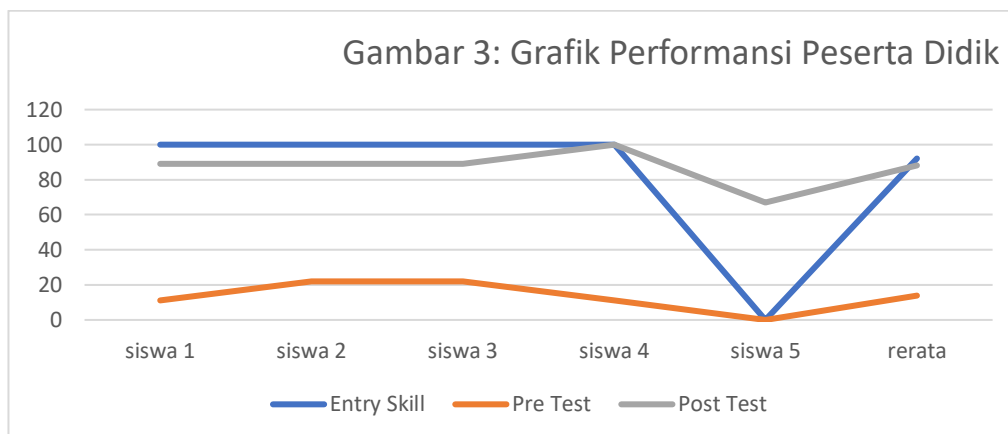
Tampak sekali dalam PPRI ini penyelenggaraan penilaian formatif dan sumatif masih bersifat umum. Sementara itu dalam Permendikbud Nomor 23 Tahun 2016 tentang Standar Penilaian Pendidikan diatur pada pasal 1 ayat (1) bahwa "Standar Penilaian Pendidikan adalah kriteria mengenai lingkup, tujuan, manfaat, prinsip, mekanisme, prosedur, dan instrumen penilaian hasil belajar peserta didik yang digunakan sebagai dasar dalam penilaian hasil belajar peserta didik pada pendidikan dasar dan pendidikan menengah" (Menteri, 2016). Tampak bahwa kecenderungan dari Permendikbud RI ini hanya mengatur tentang penilaian hasil belajar dengan menggunakan evaluasi formatif dan sumatif sebagaimana tersirat dalam Pasal 6 ayat (2) Penilaian hasil belajar oleh pendidik digunakan untuk:

- a. mengukur dan mengetahui pencapaian kompetensi Peserta Didik;
- b. memperbaiki proses pembelajaran; dan
- c. menyusun laporan kemajuan hasil belajar harian, tengah semester, akhir semester, akhir tahun. dan/atau kenaikan kelas (Menteri, 2016).

Kecenderungan standar penilaian pendidikan hanya pada hasil belajar peserta didik, dikarenakan masih banyak peraturan perundang-undangan lainnya yang mengatur evaluasi program oleh pemerintah misalnya oleh Badan Akreditasi Nasional. Implementasi Scriven Model ini lebih terlihat saat pendidik melakukan pengembangan perangkat pembelajaran dengan model pengembangan tertentu. Misalnya dalam pengembangan menggunakan model perancangan sistematis Dick&Carey, diperoleh data kuantitatif seperti berikut saat uji coba terbatas.

Tabel 1: Hasil Tes Keterampilan Masukan, Tes Awal, Tes Akhir

Student Number	3 Entry Skill Objectives	9 Pretest Instructional Objectives	9 Posttest Objectives
1	100	11	89
2	100	22	89
3	100	22	89
4	100	11	100
//			
20	67	0	67
Mean	92	14	88



Berdasarkan data kuantitatif saat evaluasi formatif uji kelompok terbatas, diperoleh gambaran bahwa program yang diimplementasikan menunjukkan keberhasilan yang signifikan. Terlihat semua subyek uji coba mengalami kenaikan hasil belajar yang signifikan.

Sedangkan untuk evaluasi sumatif, model Dick & Carey menekankan pada analisis kongruensi dan analisis dampak, yang salah satu contoh hasilnya seperti berikut.

Tabel 2: Format Ringkasan Informasi Analisis Kongruensi

Instruction reviewed _____			
Reviewer _____			
Circle level of congruence observed			
1 Not congruent,	2 Some congruence	3 Mostly congruent,	4 Right on
Statements of Organizations Characteristics	Instructional Material	Level of Congruence	Reviewer's Comments
Organization's instructional needs (goals and main objectives)	Stated goals and objectives in materials	1 2 3 4	
...			
Organization's resources available for implementing instruction	Costs within budgeted amounts	1 2 3 4	

Dimodifikasi dari (Dick, Carey, & Carey, 2015)

Implementasi Scriven Model bervariasi cara menyajikan, menganalisis, dan menginterpretasi data dan informasi. Perbedaan tersebut dapat terjadi tergantung pada tujuan yang menugaskan evaluator pelaksanaannya. Posisi evaluator ada yang berasal dari internal tim pengembang, namun banyak juga evaluator dari eksternal.

C. Tyler Model

1. Konsepsi Tyler Model

Ralph Tyler terlahir di Chicago pada tahun 1902, semula seorang guru yang kemudian diangkat sebagai Pengarah di Lembaga Penelitian Pendidikan pada *University Ohio State*, setelah lulus program doctor (PhD.). Ralph Tyler menggunakan istilah "evaluasi pendidikan" kembali pada 1930-an, membuatnya menjadi salah satu sarjana paling awal di bidang ini. (Glatthorn, Boschee, Whitehead, & Boschee, 2019, p. 502).

Model *black box* berdasarkan dua hal pokok yang menjadi fokus dari evalusinya. *Pertama*, evaluasi yang dilakukan dengan tujuan pada tingkah laku peserta didik. *Kedua*, evaluasi harus dilakukan pada awal tingkah laku peserta didik sebelum dilaksanakannya kurikulum dan sesudah melaksanakan kurikulum (hasil). Proses tidak diperhatikan dalam model Tyler ini, sehingga disebut sebagai *Black-Box Model*. Tyler juga mengatakan bahwa perubahan perilaku itu terjadi karena kegiatan kurikulum, bukan oleh pengalaman belajar. Dari kedua hal tersebut model Tyler lebih

menitik beratkan kepada hasil belajar atau hasil akhir dari pelaksanaan kurikulum, sehingga aspek proses tidak diperhatikan dalam menentukan perubahan tingkah laku atau penentuan hasil belajar. *Goal Oriented Evaluation (GOE) Model* nama lain dari *Black Box Model* merupakan sebuah model evaluasi program yang berorientasi pada tujuan. Sejak awal kegiatan dan berlangsung secara berkesinambungan GOE Model selalu menekankan pada tujuan. Penggunaan model Tyler ini membutuhkan informasi perubahan perilaku, terutama sebelum dan sesudah terjadinya implementasi program.

Model Tyler mempunyai pendekatan terhadap evaluasi terutama terdiri dari menetapkan tujuan pendidikan dan kemudian menentukan apakah tujuan tersebut telah terpenuhi. Seorang evaluator bertemu dengan pendidik untuk menentukan tujuan yang luas dan perilaku siswa yang diinginkan, yang para guru berharap untuk melihat dalam pembelajaran berikut tujuan, sekarang lebih dikenal sebagai "hasil siswa". Kemudian para pendidik seharusnya merancang kurikulum untuk mengajarkan apa yang dibutuhkan untuk mencapai tujuan. Evaluator memberikan saran tentang pengembangan langkah-langkah yang menentukan agar tujuan tercapai. Hasil penilaian dibandingkan dengan hasil yang diinginkan untuk menilai tentang efektivitas pembelajaran. Tyler mungkin paling dikenal karena studi evaluasinya yang dikenal sebagai Studi Delapan Tahun, yang melibatkan evaluasi efektivitas inisiatif pendidikan di seluruh negeri (Smith & Tyler, 1942). Meskipun ia tidak menggunakan kelompok eksperimental dan kontrol, ia mengemukakan bahwa pembentukan tujuan yang jelas dan pengukuran hasil yang ketat adalah komponen kunci dari evaluasi pendidikan (Mertens & Wilson, 2019, p. 59).

Model Tyler atau GOE Model, dapat diidentifikasi mempunyai tiga langkah pokok: (1) menentukan tujuan, (2) menciptakan situasi pencapaian tujuan, dan (3) mengembangkan alat/instrument evaluasi. Ketiga aspek ini akan menjadi perhatian saat implementasi model Tyler. Prosedur secara detail diuraikan pada prosedur Tyler Model berikut ini.

2. Prosedur Tyler Model

Tyler menyebutkan bahwa penilaian pendidikan sebagai sebuah proses untuk menentukan sejauh mana tujuan-tujuan pendidikan dari program sekolah atau kurikulum tercapai. Evaluasi berorientasi program dari Tyler ini didesain untuk menggambarkan sejauh mana tujuan program telah dicapai. Tyler menggunakan kesenjangan antara apa yang diharapkan dan apa yang berhasil diamati untuk memberikan masukan terhadap kekurangan dari suatu program. Pendekatan ini memfokuskan pada tujuan spesifik dari program dan sejauh mana program ini telah berhasil mencapai tujuan tersebut. Dalam bidang pendidikan, kegiatan yang bisa dievaluasi oleh pendekatan ini bisa saja sesederhana kegiatan-kegiatan harian di kelas atau bahkan kegiatan kompleks yang melibatkan lembaga sekolah. Hasil yang diperoleh dari evaluasi ini nantinya dapat dipakai untuk merumuskan kembali tujuan dari kegiatan, mendefinisikan kembali kegiatan/program, prosedur penilaian dan perangkat yang digunakan untuk menilai pencapaian tujuan.

Model Tyler bergerak secara rasional dan sistematis melalui beberapa langkah berikut ini.

- (1) Merumuskan tujuan perilaku, harus menentukan isi pembelajaran dan perilaku yang diharapkan.
- (2) Mengidentifikasi kondisi yang mendorong terwujudnya perilaku yang diinginkan.
- (3) Pilih, modifikasi, atau susun instrumen evaluasi yang sesuai, dan periksa instrumen untuk objektivitas, keandalan, dan validitas.
- (4) Gunakan instrumen untuk mendapatkan hasil yang diringkas atau dinilai.
- (5) Membandingkan hasil yang diperoleh dari beberapa instrumen dalam periode tertentu.
- (6) Menganalisis hasil untuk menentukan kekuatan dan kelemahan kurikulum.

(7) Menggunakan hasil sebagai dasar untuk memodifikasi yang kurikulum (Glatthorn, Boschee, Whitehead, & Boschee, 2019, p. 273).

Ketujuh langkah ini menggambarkan bahwa model GOE membatasi evaluasi program hanya pada ketercapaian tujuan dibandingkan dengan rumusan tujuan dalam program. Hasil dari model GOE ini kemudian menjadi rekomendasi bagi pengembangan program lebih lanjut.

3. Implementasi Tyler Model

Tyler model atau GOE model diimplementasikan dalam pendidikan dasar oleh pendidik hampir setiap hari untuk mengevaluasi capaian pembelajaran yang dilaksanakan. Program yang dinilai dalam pendidikan SD ini berupa Rancangan Pelaksanaan Pembelajaran Tematik. Langkah-langkah implementasi model GOE yang dilaksanakan guru profesional di SD seperti berikut.

(1) Merumuskan tujuan pembelajaran berdasarkan analisis Kompetensi Dasar yakni Indikator Pencapaian Kompetensi (IPK). Rumusan tujuan mengandung 4 (empat) unsur yaitu ABCD. Untuk evaluasi program Tyler, yang diperhatikan Behavior (perilaku) dan Degree (derajat atau standar ruang lingkup yang harus dikuasai). Contoh analisis KD:

Kompetensi Dasar	Indikator Pencapaian Kompetensi
Bahasa Indonesia 3.9 Merinci kosakata dan ungkapan pengenalan diri, keluarga, dan orang-orang di tempat tinggalnya secara lisan dan tulis yang dapat dibantu dengan kosakata bahasa daerah	3.9.1 Menganalisis kosakata anggota keluarga besar. (C4) 3.9.2 Menyusun kosakata tentang anggota keluarga besar. (C3)
4.9 Menggunakan kosakata dan ungkapan yang tepat untuk pengenalan diri, keluarga, dan orang-orang di tempat tinggalnya secara sederhana dalam bentuk lisan dan tulis.	4.9.3 Menyajikan kosa kata anggota keluarga keluarga besar secara lisan. (P3)

Tujuan pembelajaran:

1. Setelah mengamati video tentang anggota keluarga besar, peserta didik mampu mengidentifikasi sekurang-kurangnya 4 kosa kata tentang keluarga besar.
2. Setelah menyimak media power point berisi teks mengenai keluarga besar, peserta didik mampu menjelaskan sekurang-kurangnya 2 macam kegiatan bersama keluarga besar.
3. Setelah mengamati media dari media power point berisi teks mengenai keluarga besar, peserta didik mampu menuliskan kosa kata tentang anggota keluarga besar.

(2) Berdasarkan rumusan tujuan pembelajaran, dapat diidentifikasi bahwa untuk tujuan yang pertama peserta didik diberikan situasi pembelajaran audio-visual (lisan-tulisan) melalui video. Sedangkan untuk tujuan pembelajaran kedua dan ketiga, peserta didik diberikan situasi teks tertulis. Ini Sudah sesuai tuntutan perilaku yang harus dikuasai peserta didik.

(3) Berdasarkan tujuan yang dirumuskan yang sudah diberikan situasi pembelajaran yang sesuai, maka dilakukan pembuatan kisi-kisi butir soal untuk memudahkan pembuatan dan pengecekan instrument evaluasi.

KISI-KISI SOAL

Satuan Pendidikan: Sekolah Dasar
Kelas/Semester : I/1
Tema : 4. Keluargaku
Subtema : 4.3. Keluarga Besarku
Pembelajaran ke : 3

Muatan	No. KD	No. IPK	No. Tujuan	Indikator soal	Butir soal	Level	Kunci jawaban
Bahasa Indonesia	3.9	3.9.1	1	Disajikan gambar tentang anggota keluarga besar, peserta didik mampu menentukan kosa kata tentang keluarga yang benar berdasarkan gambar.	1. Amati gambar dan bacalah teks di bawah ini! (gambar kakek dan nenek) Udin memiliki keluarga besar. Udin menghormati anggota keluarganya. Gambar di samping adalah foto anggota keluarga besar Udin. Anggota keluarga yang ditandai nomor 1 dan 2 adalah a. ayah dan ibu b. paman dan bibi c. kakek dan nenek	C4	c
		3.9.2	2	Disajikan teks tentang anggota keluarga, peserta didik mampu menentukan kosa kata tentang keluarga yang benar berdasarkan gambar.	2. Udin memiliki orang tua orang tua yang penuh kasih sayang. Udin selalu bersyukur dan menghormati kedua orang tuanya. Sebutan untuk kedua orang tua Udin adalah ... a. tetangga b. ayah dan ibu c. teman dan guru	C2	b
	4.9	4.9.3	3	Disajikan 4 kosa kata secara acak, peserta didik mampu mengurutkan anggota keluarga besar dengan benar.	3. Ibu kakek nenek ayah Berdasarkan kosa kata acak di atas, Susunan keluarga besar yang benar adalah a. ayah, kakek, nenek, ibu	C4	c

- b. nenek, ibu,
- kakek, ayah
- c. kakek, nenek,
- ayah, ibu

- (4) Dibuat instrument evaluasi sesuai kisi-kisi yang sudah divalidasi oleh Subject Matter Expert (SME).
- (5) Membandingkan hasil dari instrument yang lain.
- (6) Menganalisis hasil evaluasi terhadap capaian tujuan pembelajaran siswa.
- (7) Memberikan rekomendasi terhadap kelayakan ketercapaian tujuan pembelajaran.

D. GFE Model

1. Konsepsi GFE

Mengetahui tujuan awal dari program menyebabkan terjadinya bias evaluasi program, dengan bebas tujuan dapat meningkatkan pengetahuan tentang dampak program yang sesungguhnya (Glatthorn, Boschee, Whitehead, & Boschee, 2019, p. 504). Kontribusi utama Scriven adalah untuk mengalihkan perhatian evaluator dan administrator tentang pentingnya efek yang tidak diinginkan - pengalihan yang tampaknya sangat berguna dalam pendidikan. Jika program matematika mencapai tujuannya untuk meningkatkan keterampilan komputasi tetapi memiliki efek yang tidak diinginkan dari berkurangnya minat dalam matematika, maka itu tidak dapat dinilai benar-benar berhasil. Penekanan Scriven pada metode kualitatif juga tampaknya datang pada saat yang tepat, ketika ada ketidakpuasan yang meningkat dalam komunitas penelitian dengan dominasi metodologi kuantitatif. Seperti yang dicatat Scriven sendiri, bagaimanapun, evaluasi bebas tujuan harus digunakan untuk melengkapi, bukan menggantikan, penilaian berbasis tujuan. Digunakan sendiri, tidak dapat memberikan informasi yang cukup untuk pengambil keputusan. Beberapa kritikus telah menyalahkan Scriven karena tidak memberikan arahan yang lebih eksplisit untuk mengembangkan dan menerapkan model bebas tujuan, Dan berpendapat bahwa itu mungkin hanya dapat digunakan oleh para ahli yang tidak memerlukan panduan eksplisit dalam menilai kebutuhan dan mendeteksi efek.

Dalam *Goal Free Evaluation*, Scriven mengemukakan bahwa dalam melaksanakan evaluasi program evaluator tidak perlu memperhatikan apa yang menjadi tujuan program. Yang perlu diperhatikan dalam program tersebut adalah bagaimana kerjanya (kinerja) suatu program, dengan jalan mengidentifikasi penampilan-penampilan yang terjadi (pengaruh) baik hal-hal yang positif (yaitu hal yang diharapkan) maupun hal-hal yang negatif (yang tidak diharapkan).

Asumsi yang dibangun untuk evaluasi program bebas tujuan ini ada 3 (tiga) yang mempengaruhi realisasi tujuan program. Pertama, pengaruh *side-effect* negative yang tidak dikehendaki karena akan membuat bias hasil evaluasi program. Kedua, pengaruh positif sesuai tujuan yang dikehendaki. Ketiga, pengaruh efek samping positif di luar tujuan yang dikehendaki.

Ciri-ciri evaluasi bebas tujuan atau GFE model yaitu:

- (1) Menghindari mengetahui tujuan program.
- (2) Tujuan yang telah dirumuskan terlebih dahulu tidak dibenarkan menyempitkan fokus evaluasi.
- (3) Evaluasi bebas tujuan berfokus pada hasil yang sebenarnya, bukan pada hasil yang direncanakan.
- (4) Hubungan evaluator dan manajer atau dengan karyawan proyek dibuat seminimal mungkin.
- (5) Evaluasi menambah kemungkinan ditemukannya dampak yang tidak diramalkan

Scriven menekankan evaluasi merupakan interpretasi keputusan ataupun eksplanasinya. Evaluator merupakan pembuat keputusan dan sekaligus penyedia

informasi. Dengan demikian ia membedakan antara “*Goal of evaluation* dengan *role of evaluation*”. Dengan demikian, dapat ditegaskan bahwa model GFE ini hanya memperhatikan tujuan umum program, tidak mementingkan tujuan khusus. Evaluator hanya memperhatikan dan tertuju pada penampilan komponen-komponen program yang bersangkutan.

2. Prosedur GFE Model

Prosedur yang dapat ditempuh oleh seorang evaluator untuk melakukan GFE Model mengikuti langkah-langkah berikut.

(1) Mempelajari cetak biru program yang akan dievaluasi. Untuk memahami program yang akan dievaluasi seorang evaluator GFE model harus membaca dan mengkaji cetak biru program agar tidak salah arah dalam mengevaluasi.

(2) Mengidentifikasi tujuan evaluasi program. Pengaruh sampingan negative ataupun positif, terjadi jika evaluator dapat menemukan dampak sampingan negative ataupun positif yang ditimbulkan selama evaluasi implementasi program. Pengaruh positif implementasi program yang menjadi perhatian utama.

(3) Mengembangkan desain dan instrument evaluasi. Sesuai dengan kajian terhadap cetak biru program dan tujuan evaluasi, evaluator mengembangkan desain evaluasi program beserta instrumennya. Desain evaluasi program berisikan langkah-langkah dan sasaran evaluasi program, contohnya dalam mengevaluasi program pembelajaran tematik tentunya dimulai dengan menelisik analisis KD yang dilakukan perancang pembelajaran.

(4) Memastikan pelaksanaan program telah mencapai tujuannya. Suatu program yang sudah dinyatakan selesai oleh pelaksanaannya dapat mulai dilakukan evaluasi program tersebut, GFE model tidak dilaksanakan sebelum ataupun selama program dilakukan.

(5) Menjaring dan menganalisis data. Mengumpulkan data kuantitatif dan kualitatif yang dibutuhkan sesuai desain evaluasi, yang kemudian dianalisis untuk menentukan kelayakan program.

(6) Menyusun laporan evaluasi berdasarkan hasil evaluasi. Sebagaimana model evaluasi program yang lain, maka selesai evaluasi program disusun laporan yang berisikan data dan informasi yang diperoleh dari evaluasi disertai analisis dan rekomendasi dari evaluator.

(7) Pemanfaatan hasil evaluasi program. Pemanfaatan hasil evaluasi program GFE model, setidaknya dapat digunakan untuk memperbaiki kinerja program yang dievaluasi dengan merevisi bagian-bagian program yang direkomendasikan untuk direvisi oleh evaluator.

3. Implementasi GFE Model

Implementasi GFE model sangat tergantung kepada jenis programnya, untuk program pemrosesan seperti program PPG dalam jabatan berbasis jaringan yang dilaksanakan 4 (empat) angkatan selama tahun anggaran 2021. Evaluasi program GFE model baru bisa diimplementasikan untuk angkatan ke I, angkatan ke II, dan angkatan ke III yang sudah diumumkan hasilnya. Sedangkan untuk angkatan ke IV belum bisa dilaksanakan evaluasi program GFE model, karena program PPG Dalam Jabatan Angkatan ke IV masih berlangsung.

a. Rancangan PPG Dalam Jabatan berbasis jaringan meliputi beberapa kegiatan pembelajaran berikut:

(1) Informasi awal kepada peserta:

Salah satu tahapan di kegiatan PPG Dalam Jabatan adalah Pendalaman Materi. Untuk menyiapkan kegiatan tersebut, mahasiswa PPG menjalani lebih dulu kegiatan Pembelajaran Mandiri yang dijadwalkan pada tanggal 28 Juni – 14 Juli 2021 (jadwalnya sama antara PPG Kemdikbud angkatan 3 dan PPG Kemenag angkatan 1).

Ada 2 macam materi: materi Bidang Studi (Profesional) dan materi Pedagogik. Dalam belajar mandiri ini, Bapak/ibu mempelajari materi Bidang Studi

(ada 6 modul) dan mengerjakan LK (Lembar Kerja) untuk masing-masing modul materi Bidang Studi. Mahasiswa PPG belum perlu mengerjakan LK untuk materi Pedagogik karena pengerjaannya akan dilakukan setelah menggunakan LMS (Learning Management System).

Untuk PPG Kemdikbud, materi akan tersedia di SIMPKB (bisa masuk melalui <https://ppg.simpkb.id/>) pada card Materi Pembelajaran Mandiri. Selain itu materi juga bisa diakses melalui drive https://drive.google.com/drive/folders/147s465rLGjj8f48qOL_RsV-U7dvl-xvd.

File LK hanya tersedia sebanyak 1 file, bisa dicopy untuk setiap modul yang ada. Keenam LK ini diunggah pada awal pembelajaran Pendalaman Materi melalui LMS PPG. LMS digunakan setelah orientasi akademik, informasi terkait LMS menyusul.

Silahkan mempelajarinya secara mandiri/kelompok pada bidang yang sama melalui grup mapel yang sudah dibentuk. Pembelajaran mandiri bersifat tidak terbimbing (belum melibatkan dosen) dan belum menggunakan LMS.

Petunjuk cara mengerjakan LK Belajar Mandiri dapat dilihat di <https://www.youtube.com/watch?v=9aU1JvFDkeo>

(2) Evaluasi Cetak biru tahap awal pelaksanaan PPG dalam jabatan Berbasis Jaringan.

Berdasarkan informasi yang diperoleh dari 1 (satu) kelas yang terdiri dari 35 orang peserta, penyebarannya dari Propinsi Papua Barat 1 orang, Bali 4 orang, Pulau Sulawesi 6 orang, Kalimantan 4 orang, Batam 3 orang, Natuna 1 orang, Aceh 1 orang, Sumatera Utara 3 orang, Sumatera Barat 1 orang, Sumatera Selatan 2 orang, Jakarta 2 orang, Banten 1 orang, Jawa Barat 2, Jawa Tengah 2 orang, dan Jawa Timur 2 orang. Sebaran yang mencakup dari ujung barat sampai ujung timur Indonesia. Hampir semua terjangkau jaringan dengan baik, hanya 3 orang yang mengalami kendala jaringan yakni 2 orang dari Sulawesi dan 1 orang dari Jakarta.

Kondisi ini dapat diidentifikasi (1) pengaruh sampingan negatifnya adalah perbedaan waktu WIB, WITA, dan WIT menyebabkan saat dilakukan pembelajaran sinkronus dengan standar WIB diikuti terlalu malam oleh yang WITA dan WIT, pembelajaran mandiri saat *on the job training* akan terganggu kinerja tugas jabatan dan tugas kemahasiswaan; (2) peningkatan profesionalitas guru merupakan pengaruh positif yang diharapkan, dan (3) pengaruh sampingan positif berupa naiknya literasi digital, literasi kebahasaan, dan literasi finansial.

(3) Berdasarkan data dan informasi yang ada, dirancang evaluasi program berbasis jaringan.

Angket disebar dengan menggunakan format google.form (asinkronus) dan wawancara dilaksanakan melalui webmeet sinkronus.

(4) Evaluasi program akan dilaksanakan pada tanggal 3-8 Januari 2022 saat dipastikan sudah diumumkan hasil UP dan kelulusan program PPG Guru Kelas SD dalam jabatan Berbasis Jaringan angkatan ke III.

KESIMPULAN

Berdasarkan pembahasan tentang Scriven Model, Tyler Model, dan Goal Free Evaluation Model, dapat disimpulkan hal-hal berikut.

(1) Scriven Model atau disebut juga dengan Formative-Sumative Evaluation Model, merupakan model evaluasi program yang tersurat dalam peraturan perundangan dari PPRI, Permendikbud, dan juga peraturan pemerintah daerah, maupun sistem persekolahan. Dengan demikian, Scriven Model dapat diterapkan pada level makro, meso, maupun mikro. Pada level makro ada badan atau lembaga penjaminan mutu tingkat nasional, level meso ada lembaga penjaminan mutu propinsi dan kota/kabupaten (LPMP), level sekolah ada lembaga penjaminan mutu sekolah. Setiap lembaga penjaminan mutu tersebut, mempunyai tugas pokok dan fungsi sebagai penyelenggara evaluasi formatif dan

evaluasi sumatif terhadap program-program yang dijalankan oleh pemerintah pusat, pemerintah daerah, maupun sistem persekolahan.

(2) Tyler model juga masih dipakai dalam evaluasi berupa hasil belajar

(3) Model evaluasi program yang tepat untuk suatu jenis program tertentu sangat menentukan keberhasilan evaluasi program yang dievaluasi. Ada 3 (tiga) jenis program, yaitu program pemrosesan, program layanan, dan program umum. Berikut ini tabel kesesuaian model dengan jenis program.

Tabel 3: Ketepatan model evaluasi program dan jenis program

MODEL EVALUASI PROGRAM	JENIS PROGRAM		
	PEMROSESAN	LAYANAN	UMUM
Scriven Model	Tepat	Tepat	Tepat
Tyler Model	Tepat	Tepat	Tepat
GFE Model	Tepat	Tepat	Tepat
Stake's Model	Tepat	Tepat	Tepat
Kirkpatrick Model	Tepat	Tepat	Tepat
Discrepancy Model	Tepat	Tepat	Tepat
CIPP Model	Tepat	Tidak Tepat	Tepat

(4) Kemampuan evaluator menguasai model evaluasi program dan jenis program yang dievaluasi sangat berguna dan dapat meningkatkan kebermaknaan evaluasi program yang dilakukan.

DAFTAR RUJUKAN

- Arikunto, S. P., & Jabar, C. S. (2010). *Evaluasi Program Pendidikan*. Jakarta: PT Bumi Aksara.
- Dick, W., Carey, L., & Carey, J. O. (2015). *The Systematic Design of Instruction: eighth edition*. Boston: Pearson.
- Fitzpatrick, J. L., Sanders, J. R., & Worthen, B. R. (2011). *Program Evaluation Alternative Approaches and Practical Guidelines FOURTH EDITION*. Boston: Pearson.
- Glatthorn, A. A., Boschee, F., Whitehead, B. M., & Boschee, B. F. (2019). *Curriculum Leadership Strategies for Development and Implementation Fifth Edition*. Thousand Oak : SAGE Publications, Inc.
- Menteri, P. d. (2016, Juni 6). Peraturan Menteri Pendidikan dan Kebudayaan RI, Nomor 16 Tahun 2016 tentang Standar Penilaian Pendidikan. *Standar Penilaian Pendidikan*. Jakarta, Jakarta, Republik Indonesia: Dirjen Peraturan Perundang-undangan, Kemenkumham RI.
- Mertens, D. M., & Wilson, A. T. (2019). *PROGRAM EVALUATION THEORY AND PRACTICE SECOND EDITION: A Comprehensive Guide*. New York: The Guilford Press.
- Ornstein, A. C., & Hunkins, F. P. (2018). *Curriculum Foundations, Principles, and Issues: Seventh Edition*. Harlow: Pearson Education Limited.
- Panjaitan, B. P. (2009). *Evaluasi Program Pendidikan*. Medan: Penerbit Poda.
- Presiden, R. I. (2021, Maret 30). PPRI Nomor 57 Tahun 2021 Tentang Standar Nasional Pendidikan. *Standar Nasional Pendidikan*. Jakarta, Jakarta, Republik Indonesia: Kementerian Sekretariat Negara.