

## MODEL-MODEL *ALIGNMENT* ANTARA PENILAIAN DAN KURIKULUM DALAM PEMBELAJARAN MATEMATIKA

Kana Hidayati<sup>1</sup>, Elly Arliani<sup>2</sup>

<sup>1,2</sup>Jurusan Pendidikan Matematika FMIPA UNY

<sup>1</sup>kana\_hidayati@yahoo.com, <sup>2</sup>arliani\_elly@yahoo.com

### Abstrak

Berbagai perubahan kebijakan guna meningkatkan kualitas pendidikan di Indonesia terus dilakukan pemerintah. Perubahan kebijakan tersebut diantaranya adalah perubahan kurikulum yang diberlakukan di sekolah saat ini. Terkait dengan perubahan kurikulum tersebut, salah satu komponen penting yang harus diperhatikan adalah penilaian hasil belajar siswa. Penilaian merupakan salah satu komponen penting dalam sistem pendidikan karena penilaian dapat berfungsi selain untuk memantau kualitas belajar siswa juga dapat digunakan untuk tujuan akuntabilitas. Penilaian hasil belajar siswa yang akurat dan dapat dicapai hanya jika ada kesesuaian atau kesejajaran (*alignment*) antara kurikulum, apa yang dipelajari siswa, dan apa yang muncul dari siswa pada penilaian. Oleh sebab itu perlu untuk memastikan bahwa ada kesesuaian atau kesejajaran (*alignment*) antara penilaian dan kurikulum dalam rangka memperoleh kesimpulan yang valid dari hasil penilaian.

Studi kesejajaran dalam pendidikan dimaksudkan untuk menunjukkan sejauh mana penilaian yang dilakukan mencerminkan standar isi yang harus dicapai. Hasil dari studi kesejajaran dapat digunakan juga sebagai bukti validitas untuk mendukung interpretasi skor siswa. Selama lebih dari satu dekade, berbagai metode atau model *alignment* untuk mengevaluasi kesejajaran antara penilaian dan kurikulum terus berkembang. Namun berbagai model *alignment* tersebut selama ini secara spesifik masih jarang dikaji atau diterapkan dalam kegiatan pengukuran pendidikan khususnya di Indonesia. Melalui studi literatur, artikel ini membahas berbagai model *alignment* yang dapat digunakan dalam kegiatan pengukuran pendidikan yakni untuk mengevaluasi kesejajaran antara penilaian dan kurikulum khususnya dalam pembelajaran matematika.

**Kata Kunci:** Model *Alignment* antara Penilaian dan Kurikulum, Matematika

### A. PENDAHULUAN

Reformasi bidang pendidikan di Indonesia terus ditingkatkan pemerintah karena pendidikan bagi rakyat Indonesia merupakan program penting yang sangat mendasar bagi kemajuan bangsa Indonesia di masa yang akan datang. Berbagai perombakan sistem pendidikan terus dilakukan dan disosialisasikan. Perombakan tersebut diantaranya adalah pengembangan kurikulum. Hal ini mengingat bahwa salah satu kunci untuk menentukan kualitas lulusan adalah kurikulum.

Setiap kurun waktu tertentu, kurikulum selalu dievaluasi untuk kemudian disesuaikan dengan perkembangan ilmu pengetahuan, kemajuan teknologi, dan kebutuhan pasar. Selain itu, dalam proses pengendalian mutu, kurikulum merupakan perangkat yang sangat penting karena menjadi dasar untuk menjamin tercapainya kompetensi yang diharapkan. Sejak tahun 1945 hingga saat ini, kurikulum pendidikan nasional telah mengalami perubahan, yaitu pada tahun

1947, 1952, 1964, 1968, 1975, 1984, 1994, 2004, 2006, dan 2013. Seiring dengan berubahnya kurikulum yang berlaku, sistem penilaiannya pun tentu saja juga mengalami perubahan.

Penilaian merupakan salah satu komponen penting dari sistem pendidikan karena penilaian dapat berfungsi untuk memantau kualitas belajar siswa dan untuk tujuan akuntabilitas. Penilaian hasil belajar siswa yang akurat dapat dicapai hanya jika ada kesesuaian antara kurikulum, apa yang dipelajari siswa pelajari, dan apa yang muncul dari siswa pada penilaian. Oleh karena itu perlu untuk memastikan bahwa ada kesesuaian atau kesejajaran antara kurikulum dan penilaian dalam rangka memperoleh kesimpulan yang valid dari hasil penilaian. Penilaian seharusnya memberikan informasi tentang seberapa baik siswa telah mencapai pengetahuan dan keterampilan yang diharapkan.

Salah satu strategi yang dapat digunakan untuk mengevaluasi kesesuaian antara kurikulum dan penilaian adalah dengan melakukan uji kesejajaran. Menurut Bhola, Impara, dan Buckendahl (2003), kesejajaran merupakan tingkat kesesuaian antara standar isi yang ditetapkan pemerintah dengan penilaian yang digunakan untuk mengukur prestasi siswa. Studi kesejajaran akan menunjukkan sejauh mana penilaian yang dilakukan mencerminkan standar isi yang harus dicapai. Hasil dari studi kesejajaran dapat juga digunakan sebagai bukti validitas untuk mendukung interpretasi skor tes. Ananda (2003a) menyatakan, kesejajaran dapat menjadi sumber untuk bukti validitas isi dan konstruk. Kesejajaran bisa menjadi sumber bukti validitas isi karena berusaha untuk menetapkan sejauh mana tes mencerminkan kurikulum. Bila komponen penilaian dan kurikulum dalam pendidikan memiliki kesejajaran, maka dari proses pendidikan yang dilangsungkan diharapkan menjadi efisien dan siswa memperoleh serta mampu mencapai kemampuan sesuai dengan apa yang diharapkan (Biggs, 2003).

Selama lebih dari satu dekade, berbagai metode atau model untuk mengevaluasi kesejajaran antara penilaian dan kurikulum terus berkembang. Namun berbagai model *alignment* tersebut selama ini secara spesifik masih jarang dikaji atau diterapkan dalam kegiatan pengukuran pendidikan khususnya di Indonesia. Melalui studi literatur, artikel ini membahas berbagai model *alignment* yang dapat digunakan dalam kegiatan pengukuran pendidikan yakni untuk mengevaluasi kesejajaran antara penilaian dan kurikulum khususnya dalam pembelajaran matematika di Indonesia.

## **B. PEMBAHASAN**

### **Kesejajaran antara Penilaian dan Kurikulum**

Tujuan utama dari evaluasi kesejajaran antara penilaian dan kurikulum adalah untuk memastikan bahwa antara penilaian dan kurikulum terkoordinasi dengan baik. Hasil dari beberapa studi kesejajaran memberikan informasi tentang seberapa baik penilaian telah dilakukan sesuai dengan kurikulum yang digunakan. Selain itu, kesenjangan konten dalam penilaian dapat ditentukan (Ananda, 2003a) dan informasi tersebut penting bagi para pembuat kebijakan untuk membuat keputusan tentang penilaian dan kurikulum.

Tindal (2005) menambahkan bahwa hasil dari studi kesejajaran dapat juga digunakan untuk mengidentifikasi muatan dari standar isi yang mungkin perlu diperjelas sehingga perkembangan pengetahuan di kelas juga lebih jelas. Hasil dari studi kesejajaran juga dapat digunakan dalam menentukan apakah restrukturisasi penilaian diperlukan atau tidak. Jika restrukturisasi diperlukan, hasil kesejajaran akan membantu untuk mengidentifikasi perubahan yang diperlukan dalam penilaian. Ananda (2003b) juga menyebutkan bahwa hasil kesejajaran dapat juga digunakan untuk memberikan bukti validitas isi dari sumber eksternal. Terkait penilaian dan kurikulum, Webb (1997) menyatakan bahwa studi kesejajaran memberikan informasi tentang sejauh mana dan seberapa baik antara penilaian dan kurikulum tersebut memfasilitasi dan mampu meningkatkan hasil belajar siswa.

Penilaian yang baik semestinya sejajar dengan kurikulum yang digunakan. Hal ini mengingat bahwa kesejajaran ini penting bagi efektivitas sistem pendidikan (Webb, 1997), pembelajaran siswa (Anderson, 2002; Biggs, 2003; Farenga, Joyce & Ness, 2002; La Marca, Redfield, Musim Dingin, Bailey & Hansche, 2000), keputusan akuntabilitas (Koretz & Hamilton, 2006; La Marca, 2001), evaluasi reformasi pendidikan (Herman, Webb & Zuniga, 2007), validasi

terhadap interpretasi dari skor hasil penilaian (La Marca, 2001; Rothman, 2003), dan memberikan informasi kepada siswa, orang tua, masyarakat dan pengambil kebijakan (Herman, Webb & Zuniga, 2007). Menurut Fuhrman (2001), kesejajaran ini merupakan landasan penting yang harus dapat dipenuhi terutama dalam pendidikan berbasis standar (Fuhrman, 2001).

Berdasarkan uraian di atas, menunjukkan bahwa studi kesejajaran merupakan salah satu kegiatan penting dalam kegiatan pendidikan di Indonesia. Hal ini mengingat bahwa kurikulum yang berlaku di Indonesia merupakan kurikulum yang berbasis standar.

### **Model-Model Kesejajaran**

Berdasarkan arah pendekatannya, studi kesejajaran dapat dibedakan menjadi dua yaitu kesejajaran horizontal dan kesejajaran vertikal (Niebling et al., 2008). Kesejajaran horizontal, misalnya menguji kesejajaran pada dua komponen seperti isi kegiatan pembelajaran dengan standar kompetensi pada tingkat kelas yang sama atau menguji kesejajaran pada satu komponen misalnya isi pembelajaran pada dua guru yang berbeda. Adapun kesejajaran vertikal dimaksudkan bahwa uji kesejajaran dilakukan misalnya pada komponen penilaian pada berbagai tingkat kelas yang berbeda.

Menurut Bhola et al. (2003), model kesejajaran yang berkembang dapat dikategorikan dalam tingkatan rendah, sedang, dan tinggi terkait kompleksitasnya. Kompleksitas rendah apabila hanya fokus pada perbandingan antara butir penilaian dan standar. Sedangkan kompleksitas tinggi apabila selain membandingkan antara butir penilaian dan standar juga mempertimbangkan dimensi lain seperti kedalaman isi dan tingkat penekanan dalam kurikulum dan penilaian. Oleh sebab itu, hampir semua metode kesejajaran melibatkan ahli (pakar). Para ahli ini awalnya dilatih untuk memastikan bahwa mereka mengerti dengan jelas standar, kriteria kesejajaran, dan skala yang digunakan untuk menilai kesejajaran.

Ada lima model yang bisa digunakan untuk studi kesejajaran. Menurut Bhola et al. (2003), model kesejajaran bisa dikategorikan dalam kompleksitas rendah, sedang, dan tinggi. Kategorisasi ini didasarkan pada jumlah dimensi dipertimbangkan dalam model tersebut.

#### ***Model CBE (Council for Basic Education)***

Model CBE menggunakan empat dimensi: konten, keseimbangan konten, ketelitian, dan jenis respons butir (Bhola et al, 2003.). Dimensi konten terlihat pada perbandingan antara isi butir dan standar. Keseimbangan konten berkaitan dengan distribusi butir-butir menilai standar. Jenis respon butir mengevaluasi kesesuaian jenis respon yang dicari dari siswa dalam menilai kompetensi atau keterampilan yang ditetapkan dalam standar. Namun, model ini memiliki kelemahan diantaranya bahwa pada model ini tidak menjelaskan kriteria yang jelas untuk menilai kesejajaran.

#### ***Model Penyelarasan SEC (Survey of Enacted Curriculum)***

SEC adalah model kesejajaran dengan kompleksitas sedang. Pengembangan model ini didorong oleh kebutuhan yang dirasakan untuk mengembangkan deskriptor yang seragam dari topik dan kategori kognitif yang bersama-sama dapat menggambarkan isi dari proses pembelajaran (Porter, 2002, p. 4). Salah satu keunikan dari SEC adalah bahwa model ini tidak hanya berusaha untuk membangun kesejajaran antara kurikulum (standar) dan penilaian, tetapi juga termasuk isi dari proses pembelajaran ke dalam gambar. Dengan demikian, model keselarasan SEC memuat konten dari penilaian, standar, dan proses pembelajaran. Model SEC memiliki dua dimensi dasar yaitu perbandingan isi dan kategori kognitif, yang dinilai secara bersamaan oleh ahli. Model SEC memuat lima kategori kebutuhan kognitif yakni menghafal, melakukan prosedur, berkomunikasi, memecahkan masalah non-rutin, dan generalisasi/membuktikan.

#### ***Model La Marca***

Salah satu model kesejajaran dengan kompleksitas tinggi diusulkan oleh La Marca dan rekan-rekannya (2000). Model ini memiliki perbandingan konten secara mendalam, penekanan konten, perbandingan kinerja, dan aksesibilitas sebagai dimensi (Bhola et al., 2003). La Marca, et al. (2000) menganjurkan untuk evaluasi kesejajaran antara penilaian dan standar di luar konten

yang sederhana dengan alasan bahwa kecocokan konten dapat dianggap sebagai kondisi yang diperlukan untuk kesejajaran sistem penilaian, tetapi tidak cukup untuk menghasilkan kesejajaran tingkat tinggi saja.

Untuk model La Marca, dimensi perbandingan konten mengevaluasi kesesuaian antara isi penilaian dan konten standar. Perbandingan konten mendalam menilai tingkat kesepakatan antara kompleksitas kognitif yang digariskan dalam standar dan yang tercermin dalam penilaian. Dimensi penekanan mengevaluasi kesesuaian antara bobot yang diberikan pada daerah konten tertentu dalam penilaian dan dalam standar. Menurut La Marca et al. (2000), aksesibilitas dapat dicapai jika penilaian meliputi butir yang bervariasi dalam kesulitan guna mengungkap berbagai tingkat penguasaan di tingkat kelas tertentu. Dengan demikian penilaian harus memberi kesempatan kepada semua siswa untuk menunjukkan berbagai pengetahuan dan keterampilan. Keterbatasan utama dari model ini adalah bahwa hal itu tidak memberikan petunjuk tentang bagaimana masing-masing dimensi dapat dievaluasi. Dengan kata lain, model tersebut tidak memberikan penjelasan pedoman seperti apa tingkat kesepakatan antara penilaian dan standar yang diterima.

#### **Model Webb**

Webb (1997) mengembangkan model kesejajaran dengan lima kategori yaitu fokus konten, artikulasi lintas kelas dan usia, keadilan dan kejujuran, implikasi pedagogis, dan sistem penerapan. Setiap kategori memiliki beberapa kriteria untuk menilai kesejajaran. Namun, fokus konten adalah kategori yang telah diterapkan secara luas di sebagian besar studi kesejajaran yang menerapkan model Webb. Model kesejajaran Webb merupakan alat yang ampuh yang dapat digunakan untuk membandingkan hasil pada seluruh wilayah negara. Perbandingan ini dimungkinkan karena data kuantitatif yang dihasilkan dari model ini. Namun, hasil dari kesejajaran model Webb kadang bisa menyesatkan. Misalnya, Martone dan Sireci (2009) mencatat bahwa butir yang mengukur hanya bagian dari tujuan yang lebih luas, dinyatakan masih dianggap sesuai dengan tujuan. Dengan demikian, hasil dari kesejajaran dapat meningkat sejauh persetujuan kategoris dari ahli berbagai pengetahuan dan keseimbangan representasi yang bersangkutan.

#### **Model Achieve**

Model keselarasan *Achieve* memiliki enam criteria yaitu akurasi tes, sentralitas konten, sentralitas kinerja, tantangan, keseimbangan, dan jangkauan (Bhola et al. 2003). Proses kesejajaran model ini menggunakan mengikuti tiga tahap. Pertama adalah butir dianalisis dengan analisis butir di mana butir dibandingkan dengan standar untuk mengkonfirmasi draft tes, menilai sentralitas konten, dan mengevaluasi sentralitas kinerja. Tahap kedua, menilai tantangan dalam hal sumber dan tingkat dan tahap terakhir menilai keseimbangan dan jangkauan. Konfirmasi dari uji draft tes melibatkan ahli yang mencocokkan setiap butir draft untuk memastikan bahwa setiap butir dalam penilaian tersebut terkait dengan setidaknya satu tujuan dalam standar. Para ahli melakukan ini dengan cara diskusi untuk mencapai konsensus tentang tingkat kecocokan antara butir dan tujuan yang berkaitan. Butir ini dianggap sesuai dengan tujuan jika mengukur konten yang sama dengan yang ditentukan dalam standar (Rothman, Slattery & Vranek, 2002). Ketersediaan data kualitatif pada model *Achieve* menyediakan pemahaman menyeluruh untuk tingkat kesejajaran. Informasi ini dapat digunakan untuk meninjau kesejajaran antara penilaian dan standar. Namun, penggunaan model ini membutuhkan banyak waktu dan personal yang terampil, serta biaya yang tinggi.

Berdasarkan uraian di atas, semua model kesejajaran mengandalkan ahli untuk menilai derajat kesejajaran antara penilaian dan kurikulum (standar). Kualitas hasil kesejajaran tergantung pada seberapa baik ahli memahami kriteria penilaian selama pelatihan. Dalam hal menilai kesejajaran, semua model mengevaluasi perbandingan dalam konten antara butir dalam penilaian dan standar dalam kurikulum. Hal ini membantu untuk memeriksa bahwa setiap butir pada penilaian mengukur konten dalam beberapa tujuan. Model-model kesejajaran juga mengevaluasi sejauh mana luasnya pengetahuan dalam penilaian mencerminkan luasnya pengetahuan dalam standar. Kelima model menilai tingkat kesepakatan pada tuntutan kognitif yang ditentukan dalam standar dan yang dibutuhkan untuk ujian guna memberikan respons yang

benar untuk butir pada penilaian. Meskipun tingkat tantangan adalah aspek yang sangat penting dalam kesejajaran, semua model menggunakan hasil pembahasan ahli untuk menilai kesejajaran itu.

Sejumlah perbedaan dari beberapa model kesejajaran memberikan kriteria berbeda untuk menilai kesejajaran (misalnya, Webb dan Achieve), sementara yang lainnya tidak (misalnya, La Marca). Kurangnya kriteria untuk menilai kesejajaran membatasi utilitas dari model tersebut. Model kesejajaran juga berbeda dalam hal tingkat detail untuk pencocokan standar dalam penilaian. Dalam beberapa metode, pencocokan dilakukan pada tingkat standar yang lebih global. Model Webb adalah satu-satunya model yang dapat mengakomodasi pencocokan pada setiap tingkat standar seperti perbedaan hasil serta komparabilitas terutama jika komponen yang dievaluasi dalam studi kesejajaran (misalnya, penilaian dan standar) ditulis pada tingkat yang berbeda detail. Terkait dengan hal ini menunjukkan bahwa beberapa metode memberikan hasil kesejajaran baik kualitatif maupun kuantitatif (misalnya, Webb, SEC, dan Achieve) sementara yang lainnya tidak (misalnya, CBE dan La Marca). Hasil kuantitatif dan kualitatif penting dalam membandingkan hasil seluruh wilayah negara dan kekurangan dalam penilaian atau kurikulum. Perbedaan penting lainnya adalah bahwa hanya metode penyelarasan SEC yang menggabungkan proses pembelajaran ke dalam kesejajaran. Hal ini membantu dalam memberikan informasi di bagian dari kurikulum yang berfokus pada guru.

Berdasarkan uraian di atas, menunjukkan bahwa kriteria konten termasuk dalam kerangka taksonomi Webb (Webb, 1997) tapi ternyata secara spesifik tidak menawarkan alat untuk mengkategorikan konten. Oleh karena itu, taksonomi yang ada dalam model Webb masih memiliki kelemahan yang perlu diperhatikan khususnya terkait konten kognitifnya. Terkait dengan kompleksitas kognitif, berbagai taksonomi telah dikemukakan para ahli seperti taksonomi Bloom yang direvisi (Anderson & Krathwohl, 2001), DeBlock (de Landsheere, 1990), De Corte (de Landsheere, 1990), Guilford (1967), Marzano (2001), matriks kinerja-konten Merrill (1994); PISA (OECD, 1999), Porter (Porter & Smithson, 2001a, 2001b) dan TIMSS (Robitaille et al., 1993). Namun taksonomi Bloom yang direvisi (Anderson & Krathwohl, 2001) dalam perkembangannya ternyata juga dapat digunakan untuk mengevaluasi kesejajaran antara penilaian dan kurikulum dan cocok untuk diterapkan dalam pembelajaran matematika (Gunilla N & Henriksson W, 2008).

### C. SIMPULAN DAN SARAN

Berdasarkan hasil kajian di atas, dapat diambil simpulan bahwa model-model kesejajaran yang berkembang dan dapat digunakan dalam pendidikan diantaranya adalah Webb, SEC, Achieve, CBE, dan La Marca. Mengingat bahwa sistem pendidikan di Indonesia berbasis standar, penggunaan kelima model tersebut dapat dilakukan namun sebaiknya memperhatikan kelemahan masing-masing model disesuaikan dengan tujuan studi kesejajaran yang dilakukan. Penggunaan model webb dan taksonomi Bloom yang direvisi lebih disarankan dalam kegiatan evaluasi kesejajaran antara penilaian dan kurikulum dalam pembelajaran matematika di Indonesia mengingat kompleksitas kognitif yang dikajinya.

### D. DAFTAR PUSTAKA

- Ananda, S. 2003a. *Rethinking issues of alignment under No Child Left Behind*. San Francisco: WestEd.
- Ananda, S. 2003b. Achieving alignment. *Leadership*, 33(1), 18-21.
- Anderson, L. W. 2002. Curricular alignment: A re-examination. *Theory in Practice*, 41(4), 255-260.
- Anderson, L. W., & Krathwohl, D. R. 2001. *A taxonomy for learning, teaching, and assessing. A revision of Bloom's taxonomy of educational objectives*. New York: Addison Wesley Longman.

- 
- Bhola, D. S., Impara, J. C., & Buckendahl, C. W. 2003. Aligning Tests with content Standards: Methods and Issues. *Educational Measurement: Issues and Practice*, 22(3), 21-29.
- Gunilla N & Henriksson W. 2008. Alignment of Standards and Assessment: A Theoretical and Empirical Study of Methods for Alignment. *Electronic Journal of Research in Educational Psychology*, 6(3), 667-690
- Herman, J., Webb, N., & Zuniga, S. 2005. *Measurement issues in alignment of standards and assessment: A case study*. (CSE Report 653). Los Angeles; University of California, National Center for Research on Evaluation, Standards and Student Testing (CRESST).
- La Marca, P. M., Redfield, D., Winter, P. C., Bailey, A. & Despriet, L. 2000. *State Standards and State Assessment Systems: A guide to alignment*. Washington, DC; Council of Chief State Officers.
- Martone, A. & Sireci, S. G. 2009. Evaluating alignment between curriculum, assessments, and instruction. *Review of Educational Research*, 79(4), 1332 - 1361.
- Martone, D., Sireci, S. G., & Delton, J. 2006. *Methods for the alignment between state curriculum frameworks and state assessments: A literature review*. Center for Educational Assessment Research Report No 603. Amherst, MA: University of Massachusetts, School of Education.
- Tindal, G. 2005. *Alignment of Alternate Assessments using the Webb System*. Washington, DC; Council of Chief State Officers.
- Webb, N. L. 1997. *Criteria for Alignment of Expectations and Assessments in Mathematics and Science Education*. Research Monograph No. 6). Washington DC: Council of Chief State Officers.
- Webb, N. L., M., Ely, R., Cormier, M. & Vesperman, B. (2005). *The WEB Alignment Tool: Development, Refinement, and Dissemination*. Washington, DC; Council of Chief State Officers.
- Webb, N. L. 2006. *Alignment Analysis of Mathematics Standards and Assessments. Wisconsin, Grades 3-8 and 10*. Retrieved October 20, 2009, from <http://www.dpi.state.wi.us/oea/pdf/mathsummary06.pdf>