

**PENGEMBANGAN SOAL MATEMATIKA MODEL TIMSS DI KELAS IV
SEKOLAH DASAR**

1

Putri Cahyani Agustine
STKIP Muhammadiyah Bangka Belitung
putricahyani1990@gmail.com

ABSTRACT

Mathematics is a universal science that underlies the development of modern technology that has an essential role in various scientific disciplines and advances human thinking power. One of the achievements of Indonesian students in the field of mathematics can be seen based on the 2011 TIMSS ranking in which Indonesia ranked 38th of 42 participating countries. One of the factors that cause low TIMSS results is because the TIMSS model questions in its lesson that have moderate and high complexity and require reasoning in the completion process. Therefore, researchers were interested in developing mathematical question with TIMSS model in the fourth grade of elementary school. The development of these questions does through two stages, namely the preliminary study stage, the preparation stage and the formative evaluation stage which included self-evaluation, prototyping (expert reviews, one to one, and small group), and the test field. In this study, there were ten questions categorized as valid and practical qualitatively and quantitatively.

Keywords: TIMSS, development, and question

PENDAHULUAN

1. Latar Belakang

Matematika merupakan ilmu universal yang mendasari perkembangan teknologi modern, mempunyai peranan penting dalam berbagai disiplin ilmu dan memajukan daya pikir manusia. Matematika perlu diberikan kepada semua peserta didik mulai dari sekolah dasar untuk membekali peserta didik dengan kemampuan berpikir logis, analitis, sistematis, kritis, kreatif, dan kemampuan bekerja sama. Kompetensi tersebut diperlukan agar peserta didik dapat memiliki kemampuan memperoleh, mengolah, dan memanfaatkan informasi untuk bertahan hidup pada keadaan yang selalu berubah, tidak pasti, dan kompetitif. Oleh karena itu, setiap orang diharapkan melek matematika (*mathematical literacy*) sehingga mampu menyelesaikan masalah dan mampu menghadapi persaingan global dalam

proses pengambilan keputusan dalam penyelesaian masalah di kehidupan (Lewy,2009).

2

Salah satu prestasi siswa Indonesia dalam bidang matematika, dapat dilihat berdasarkan rangking *Trends in International Mathematics and Science Study* 2011, Indonesia menempati peringkat ke 38 dari 42 negara yang berpartisipasi. Pada tahun 2007 dimana Indonesia menempati rangking 36 dari 48 negara. TIMSS merepresentasikan rentang kemampuan siswa berdasarkan benchmark internasional menjadi 4 tingkatan standar. Keempat tingkatan tersebut yaitu standar mahir (625), standar tinggi (550), standar menengah (475) dan standar rendah (400). Sementara pencapaian siswa di Indonesia sendiri masih pada tingkatan standar rendah. Hal itu bisa terlihat dari tabel persentase hasil pencapaian siswa Indonesia di dalam *TIMSS* 2011 untuk tiap-tiap domain konten dan domain kognitif jika dibandingkan dengan negara yang lain berikut ini.

Tabel 1
Rata-rata Persentase Menjawab Benar
pada Domain Konten dan Kognitif

Negara	Bilangan	Aljabar	Geometri dan Pengukuran	Data dan Peluang	Knowing	Applying	Reasoning
Singapura	➤ 7 7 (0,9)	➤ 7 2 (1,1)	➤ 71 (1,0)	➤ 7 2 (0,9)	➤ 8 2 (0,8)	➤ 7 3 (1,0)	➤ 62 (1,1)
Korea Ref	➤ 7 7 (0,5)	➤ 7 1 (0,7)	➤ 71 (0,6)	➤ 7 5 (0,5)	➤ 8 0 (0,5)	➤ 7 3 (0,6)	➤ 65 (0,6)
Jepang	➤ 6 3 (0,7)	➤ 6 0 (0,7)	➤ 67 (0,7)	➤ 6 8 (0,6)	➤ 7 0 (0,6)	➤ 6 4 (0,6)	➤ 56 (0,7)
Malaysia	➤ 3 9 (1,3)	➤ 2 8 (0,9)	➤ 33 (1,1)	➤ 3 8 (0,9)	➤ 4 4 (1,2)	➤ 3 3 (1,0)	➤ 23 (0,9)
Thailand	➤ 3 3 (1,0)	➤ 2 7 (0,9)	➤ 29 (0,9)	➤ 3 9 (0,8)	➤ 3 8 (1,0)	➤ 3 0 (0,8)	➤ 22 (0,8)
Indonesia	➤ 2 4 (0,7)	➤ 2 2 (0,5)	➤ 24 (0,6)	➤ 2 9 (0,7)	➤ 3 7 (0,7)	➤ 2 3 (,6)	➤ 17 (0,4)
Rata-rata International	➤ 4 3 (0,1)	➤ 3 7 (0,1)	➤ 39 (0,1)	➤ 4 5 (0,1)	➤ 4 9 (0,1)	➤ 3 0 (0,1)	➤ 30 (0,1)

Sumber : Mullis (2012)

Dari tabel di atas, dapat disimpulkan bahwa domain *reasoning* yang dimiliki oleh siswa Indonesia lebih rendah dari pada domain *knowing* dan

applying. Salah satu faktor yang menyebabkan rendahnya hasil TIMSS adalah karena soal-soal yang diberikan kepada siswa di ajang internasional tersebut dengan yang biasa diberikan oleh guru berbeda. Selain itu juga, dikarenakan soal-soal model TIMSS termasuk soal-soal yang memiliki kompleksitas sedang dan tinggi, serta memerlukan penalaran dalam proses penyelesaiannya dimana soal-soal seperti itu jarang diberikan oleh guru kepada siswa di dalam proses pembelajaran matematika di Indonesia (Rizta,2013).

Menurut Mullis (2013), Soal-soal TIMSS dapat menilai kemampuan siswa untuk menunjukkan pengetahuan mereka, menerapkan apa yang mereka pelajari, memecahkan masalah, dan alasan melalui analisis dan pemikiran logis. Menurut Wardhani dan Rumiati (2011), salah satu penyebab rendahnya hasil *TIMSS* adalah siswa Indonesia pada umumnya kurang terlatih dalam menyelesaikan soal-soal dengan karakteristik seperti soal-soal pada *TIMSS* yang substansinya kontekstual, menuntut penalaran, argumentasi dan kreativitas dalam menyelesaikannya. Berdasarkan penjelasan diatas, belum ada penelitian pengembangan soal matematika model TIMSS di kelas IV sekolah dasar sehingga peneliti tertarik untuk melakukan penelitian yang berjudul “**Pengembangan Soal Matematika Model TIMSS di kelas IV Sekolah Dasar**”.

Rumusan Masalah

Bagaimana mengembangkan soal matematika model kelas IV SD Islam Az-Zahrah Palembang yang valid dan praktis?

Tujuan Penelitian

Menghasilkan soal matematika model TIMSS di kelas IV SD Islam Az-Zahrah Palembang yang valid dan praktis.

2. TINJAUAN PUSTAKA

2.1 TIMSS

TIMSS adalah studi internasional tentang kecenderungan atau arah perkembangan matematika dan sains. Studi ini diselenggarakan oleh *International Association for the Evaluation of Educational Achievement (IEA)* yaitu sebuah asosiasi internasional untuk menilai prestasi dalam pendidikan. *TIMSS* berpusat di *Lynch School of Education*, Boston College, USA. *TIMSS* bertujuan untuk mengetahui peningkatan pembelajaran matematika dan sains. *TIMSS* diselenggarakan setiap 4 tahun sekali. Pertama kali diselenggarakan pada tahun 1995, kemudian berturut-turut pada tahun 1999, 2003, 2007, 2011, dan 2015. Salah satu kegiatan *TIMSS* adalah menguji kemampuan matematika siswa kelas 4 SD (Sekolah Dasar) dan kelas 8 SMP (Sekolah Menengah Pertama).

Menurut Lie (2015), *TIMSS* memiliki beberapa kriteria yaitu *Organiser, International Association for the Evaluation of Educational Achievement (IEA; Objective, To provide important background information that can be used to improve teaching and learning in mathematics and science ; cycle4-years cycle since 1999; student sample Grade 4 and 8.*

Dalam *TIMSS 2011 Assessment framework* (Mullis, Martin, Ruddock, O'Sullivan dan Preuschoff : 2011), disebutkan bahwa dimensi konten terdiri atas empat domain, yaitu: bilangan, aljabar, geometri, data dan peluang, sementara pada *TIMSS* sebelumnya dimensi konten terdiri atas lima dimensi konten karena domain data dan peluang dipisah. Tiap dimensi konten diperinci lebih lanjut dalam beberapa topik, misalnya dimensi konten bilangan meliputi topik bilangan cacah, pecahan dan desimal, bilangan bulat, perbandingan, proporsi dan persentase (Wardhani dan Rumiati, 2011).

2.2 Teori Dasar dan Kerangka Kerja *Trends In International Mathematics and Sciences Study (TIMSS)*

Menurut Mullis (2015), *TIMSS 2015 Assessment Frame Work* ada dua dimensi penilaian yaitu : dimensi konten dan dimensi kognitif. Dimensi kognitif terdiri atas tiga dimensi yaitu mengetahui fakta dan prosedur (Pengetahuan),

menggunakan konsep dan memecahkan masalah rutin (penerapan) dan memecahkan masalah nonrutin (penalaran). Dimensi kognitif dimaknai sebagai perilaku yang diharapkan dari siswa ketika mereka berhadapan dengan dimensi matematika yang tercakup dalam dimensi konten. Dalam dimensi kognitif, pemecahan masalah merupakan fokus utama dan muncul dalam soal-soal tes yang terkait dengan hampir semua topik dalam tiap dimensi konten. Ketiga dimensi dalam dimensi kognitif merupakan perilaku yang diharapkan dari siswa ketika mereka berhadapan dengan dimensi matematika yang tercakup dalam dimensi konten.

2.3 Pengembangan Soal

Penelitian pengembangan yang juga disebut *development research* berorientasi pada pengembangan produk dimana proses pengembangannya dideskripsikan seteliti mungkin dan produk akhirnya dievaluasi. Produk yang dihasilkan dalam penelitian pengembangan ini adalah perangkat soal.

Perangkat soal yang dikembangkan memiliki tiga karakteristik yaitu valid dalam segi konten, konstruks dan bahasa, yang kemudian soal divalidasi oleh pakar dan teman sejawat. Cara ini dikenal dengan triangulasi. Triangulasi adalah suatu teknik validasi data yang memanfaatkan sesuatu yang lain di luar itu (teman sejawat dan pakar) untuk keperluan pengecekan dan sebagai pembanding/dasar merevisi instrumen penilaian (Krathwohl,1997).

2.4 Validitas dan Reliabilitas

Indikator utama dari kualitas perangkat soal yang dikembangkan terletak pada uji validitas dan reliabilitas (Kimberlin & Winterstein, 2008). Menurut Fraenkel dan Wallen (2010) uji validitas dilakukan agar instrumen yang dibuat menghasilkan informasi yang sesuai, bermakna, tepat, dan berguna dalam penarikan kesimpulan yang dibuat oleh peneliti.

Kata valid berarti tepat, sehingga perangkat soal dapat dikatakan valid apabila dapat dengan tepat mengukur apa yang mau diukur. Maka dari itu, validitas dapat dikatakan sebagai ketepatan dari perangkat soal.

Validitas dibagi menjadi tiga jenis (Moskal & Leydens, 2000), yaitu:

1. Validitas Isi (*Content Validity*)

Validitas isi adalah validitas yang ditilik dari segi isi tes itu sendiri sebagai alat ukur, yaitu sejauh mana perangkat tes sebagai alat pengukur telah dapat mewakili secara representatif keseluruhan materi atau bahan pelajaran yang akan diteskan.

2. Validitas Konstruksi (*Construct Validity*)

Validitas konstruksi dapat diartikan sebagai validitas yang dilihat dari segi susunan, kerangka atau rekaannya. Suatu perangkat soal dapat dinyatakan sebagai alat ukur yang telah memiliki validitas konstruksi, apabila perangkat soal tersebut yang ditinjau dari segi susunan, kerangka atau rekaannya telah dapat dengan tepat mencerminkan suatu konstruksi dalam teori psikologis.

3. Validitas Kriteria (*Criterion Validity*)

Validitas kriteria yaitu menentukan valid atau tidaknya suatu perangkat soal berdasarkan kriteria tertentu, yaitu kriteria internal maupun kriteria eksternal. Kriteria internal adalah tes atau instrumen itu sendiri yang menjadi kriteria, sedangkan kriteria eksternal adalah hasil ukur instrumen atau tes lain di luar instrumen itu sendiri yang menjadi kriteria. Ukuran lain yang sudah dianggap baku atau dapat dipercaya dapat pula dijadikan sebagai kriteria eksternal, dalam hal ini validitas kriteria eksternal mengambil contoh-contoh dari soal-soal TIMSS.

Validitas yang ditentukan berdasarkan kriteria internal disebut validitas internal. Validitas internal mempermasalahkan validitas butir dengan hasil ukur instrumen sebagai suatu kesatuan dan sebagai kriteria sehingga bisa disebut validitas butir soal. Validitas butir tercermin dari besarnya koefisien korelasi antara skor butir dengan skor total instrumen.

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

Jika koefisien korelasi skor butir dengan skor total lebih besar dari koefisien korelasi tabel r , maka koefisien korelasi butir signifikan dan butir dianggap valid. (Djaali & Muljono, 2008).

7

Selain diuji ketepatannya, perangkat soal juga diuji ketetapannya. Suatu perangkat soal dapat dikatakan reliabel atau tetap jika dalam beberapa kali diberikan kepada siswa, menunjukkan hasil yang hampir sama atau serupa (Long & Johnson, 2000). Uji reliabilitas dilakukan untuk mengukur ketetapan dari suatu perangkat soal. Pengujian ini dimaksudkan untuk menjamin instrumen yang digunakan merupakan sebuah instrumen yang handal, konsistensi, dan stabil, sehingga bila digunakan berkali-kali dapat menghasilkan data yang sama.

Dalam rangka menentukan apakah soal-soal tes yang dikembangkan memiliki daya keajegan mengukur atau reliabilitas yang tinggi maka peneliti menggunakan rumus Alpha sebagai berikut:

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

Keterangan: r_{11} = koefisien reliabilitas tes

n = banyaknya butir item yang dikeluarkan dalam tes

$\sum S_i^2$ = jumlah varian skor dari tiap-tiap butir item

S_i^2 = varian total

Selanjutnya dalam pemberian interpretasi terhadap koefisien reliabilitas tes pada umumnya digunakan patokan sebagai berikut:

Apabila r_{11} sama dengan atau lebih besar daripada 0,60 berarti soal-soal tes yang sedang diuji reliabilitasnya dinyatakan tinggi dan apabila r_{11} lebih kecil daripada 0,60 berarti soal-soal yang sedang diuji memiliki reliabilitas rendah (Sudijono, 2009).

3. METODOLOGI PENELITIAN

3.1 Subjek dan Waktu Penelitian

Subjek penelitian adalah siswa kelas IV SD Islam Az-Zahrah Palembang. Penelitian ini telah dilakukan pada semester genap tahun ajaran 2015-2016.

3.2 Jenis Penelitian

Penelitian ini digolongkan penelitian pengembangan atau *development research*. Penelitian pengembangan sebagai jenis penelitian yang dilakukan untuk menghasilkan soal model *TIMSS* di kelas IV SD yang valid dan praktis.

Pengembangan soal ini dilakukan melalui dua tahapan yaitu tahap *priliminary study*, yaitu tahap persiapan dan tahap *formatif evaluation* (Tessmer, 1993) yang meliputi *self evaluation*, *prototyping (expert reviews dan one to one*, dan *small group*), serta *field test*.

3.3 Prosedur Penelitian

Penelitian ini telah dilakukan dalam dua tahap, yaitu tahap *preliminary* (persiapan) dan tahap *formatif evaluation*.

3.3.1 Tahap Preliminary

3.3.1.1 Analisis

Peneliti dalam hal ini telah melakukan analisis siswa untuk mengetahui tingkat kemampuan siswa (tinggi, sedang dan rendah) yang telah dijadikan validator pada tahap *One-to-one* dan *small group*, analisis kurikulum SD untuk menentukan batasan materi yang telah dipelajari siswa dan analisis soal-soal *TIMSS* untuk memastikan kesesuaian soal dengan karakteristik *TIMSS* yang dapat mengukur kemampuan berpikir tingkat tinggi.

3.3.1.2 Pendesainan Soal

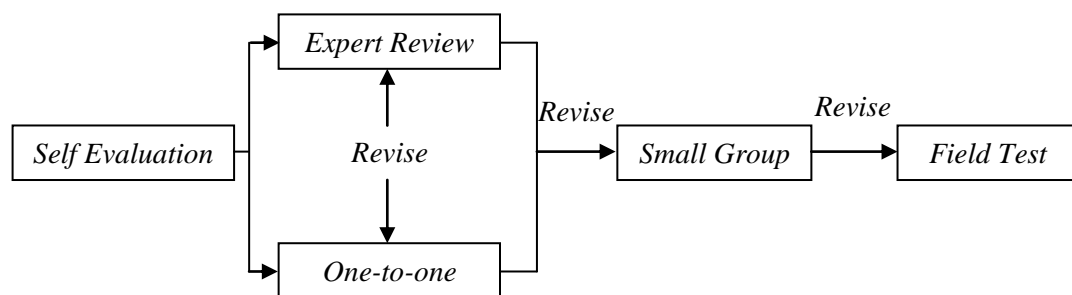
Pada tahap ini telah dilakukan pendesainan soal *TIMSS*. Pendesainan soal ini didasarkan pada tiga karakteristik, yaitu konten, konstruk, dan bahasa. Pendesainan soal ini telah menghasilkan prototipe yang telah divalidasi pada tahap *formatif evaluation*.

3.3.2 Tahap Formatif Evaluation

Pada tahap *formative evaluation* ini meliputi *self evaluation*, *expert reviews*, *one-to-one*, *small group*, dan *field test* (Tessmer, 1998; Zulkardi, 2002). Berdasarkan hal tersebut, pada penelitian ini, berikut tahapan *formative evaluation* yang telah peneliti lakukan.

Alur desain *formative evaluation* yang telah dilaksanakan pada penelitian ini dapat dilihat pada Gambar 3 berikut:

Gambar 2
Alur Penelitian Pengembangan



Berdasarkan diagram alur pengembangan bahan ajar diatas, maka peneliti mengembangkan prototipe bahan ajar melalui tahapan-tahapan sebagai berikut:

3.3.2.1 *Self Evaluation*

Pada tahap ini peneliti telah melakukan penilaian kembali terhadap desain prototipe soal TIMSS yang telah dibuat dan dari hasil penilaian tersebut telah didapatkan prototipe pertama

3.3.2.2 *Expert Reviews*

Hasil prototipe pertama diberikan kepada pakar untuk dievaluasi dari segi konten, konstruk, dan bahasa. Proses validasi dilakukan melalui email dan komentar dari validator.

3.3.2.3 *One-to-one*

Bersama dengan uji pakar (*expert review*), peneliti melakukan uji individu (*one-to-one*), dengan 3 orang siswa sebagai *tester*. Ketiga siswa tersebut diminta untuk memberikan tanggapan dan komentarnya terhadap produk yang dihasilkan, lalu hasil komentar tersebut telah digunakan untuk bahan pertimbangan revisi prototipe pertama. Hasil dari revisi *expert review* dan *one-to-one* menghasilkan

prototipe kedua.

10

3.3.2.4 *Small Group*

Prototipe kedua yang telah dihasilkan telah diujicobakan di *small group* yang terdiri dari 6 orang siswa non subjek yang memiliki karakteristik yang sama dengan karakteristik siswa yang akan dijadikan sebagai subjek penelitian. Keenam orang siswa tersebut telah diminta untuk mengerjakan soal yang telah dibuat sekaligus diminta untuk mengomentari soal tersebut. Proses pengerjaan soal telah direkam dengan menggunakan kamera video. Hasil komentar-komentar siswa dan hasil rekaman tersebut telah digunakan untuk bahan pertimbangan dalam proses revisi prototipe kedua menjadi prototipe ketiga.

3.3.2.5 *Field Test*

Hasil revisi prototipe kedua berupa prototipe ketiga telah diujicobakan ke subjek penelitian yaitu siswa kelas IV SD Islam Az-Zahrah Palembang yang dilakukan pada semester genap 2015/ 2016 dalam tahap *field test*. Prototipe yang diujicobakan pada *field test* ini merupakan prototipe terakhir yang telah memenuhi kriteria kualitas yang memiliki validitas, kepraktisan, dan efektivitas. Dari hasil *field test* ini telah dilakukan perhitungan terhadap skor dari setiap siswa dan digunakan sebagai dasar penilaian dan pengelompokan terhadap kemampuan matematis siswa.

3.4 Teknik Pengumpulan Data

Berdasarkan metode penelitian ini, teknik pengumpulan data yang telah digunakan adalah:

3.4.1 Dokumentasi

Dokumentasi berupa prototipe pertama yang diperoleh dari tahap *self Evaluation* dari peneliti sendiri, komentar *expert* melalui lembar validasi dari email yang, komentar terhadap prototipe dari tahap *one-to-one* dengan 3 orang siswa dengan kemampuan tinggi, sedang, dan rendah pada siswa kelas IV yang bukan subjek penelitian dan tahap *small group* dengan 6 orang siswa dengan kemampuan

tinggi, sedang, dan rendah pada siswa kelas IV yang bukan subjek penelitian dan hasil kerja siswa terhadap soal pada tahap *field test* dengan 23 siswa kelas IV yang merupakan subjek penelitian.

11

3.4.2 Rekaman Video

Pengumpulan data pada proses penelitian juga menggunakan rekaman video. Rekaman ini bertujuan merekam kejadian pada *one-to-one*, *small group*, dan *field test*. Rekaman video pada saat *one-to-one* terlihat masih bingung dalam pengerjaan soal ada beberapa soal yang membuat siswa bingung, pada saat *small group* siswa terlihat lebih aktif dan tidak terlihat bingung lagi dalam mengerjakan soal karena instrumen yang diberikan kepada siswa telah direvisi baik bahasa, konstruk, dan isi. Pada saat *field test* siswa terlihat lebih aktif dalam mengerjakan instrumen.

3.4.3 Tes

Tes digunakan untuk mengetahui efek potensial dari pengembangan soal matematika model TIMSS di kelas IV. Data dianalisis untuk melihat pencapaian kemampuan matematis siswa. Soal tes yang telah digunakan berjumlah 10 buah soal yang keseluruhannya berbentuk soal uraian.

3.4.4 Lembar Observasi

Lembar observasi adalah lembar kerja yang berfungsi untuk mengobservasi dan mengukur tingkat keberhasilan atau ketercapaian tujuan dalam penelitian. Lembar observasi digunakan untuk melihat kemampuan yang tampak dalam penelitian.

3.4.4 Lembar Angket

Lembar angket adalah lembar yang berfungsi untuk mengukur tingkat keberhasilan atau ketercapaian tujuan dalam penelitian.

3.5 Teknik Analisis Data

Teknik analisis data yang digunakan di dalam penelitian ini adalah analisis deskriptif kualitatif. Data yang dianalisis secara deskriptif kualitatif berasal dari dokumentasi, rekaman video dan tes tertulis.

3.5.1 Dokumentasi

Peneliti telah menganalisis prototipe pertama, menganalisis komentar yang diberikan oleh *expert review* dalam lembar validasi, menganalisis komentar yang diberikan oleh siswa pada tahap *one-to-one*, dan menganalisis hasil kerja siswa pada tahap *field test*.

3.5.2 Rekaman Video

Peneliti telah menganalisis hasil rekaman video untuk melihat kepraktisan penggunaan soal oleh siswa pada tahap *one-to-one*, *small group* dan *field test*.

3.5.4 Lembar Observasi

Lembar observasi digunakan untuk melihat pada deskriptor kemampuan siswa yang dominan tampak dalam penelitian.

3.4.4 Lembar Angket

Lembar angket digunakan untuk melihat komentar siswa terhadap soal yang telah dikerjakan dalam penelitian.

Jadi, Penelitian ini dikatakan berhasil jika telah menghasilkan soal TIMSS di kelas IV yang valid dan praktis serta dapat digunakan dan diterapkan. Kevalidan dari soal-soal secara kualitatif didapat berdasarkan dari hasil validasi pakar (*expert review*) dan *one-to-one* yaitu berupa saran, komentar dan masukan yang menyattelah bahwa soal yang telah dikembangkan oleh peneliti sudah valid baik dari segi konten, konstruk maupun bahasa. Kevalidan secara kuantitatif didapat berdasarkan analisis butir soal.

4. PEMBAHASAN

Perangkat instrumen yang dihasilkan pada tahap *self evaluation* dinamakan dengan *prototype* 1 yang merupakan soal-soal model TIMSS siswa kelas IV SD berjumlah 10 butir soal.

4.1 Prototyping

4.1.1 Self Evaluation

Hasil dari desain diperoleh *prototype* 1 yang merupakan soal model TIMSS tipe berpikir tingkat tinggi untuk siswa kelas IV SD berjumlah 10 soal, kemudian diperiksa oleh peneliti. Jika di dalam proses pendesainan terdapat kesalahan

pengetikan, kesalahan pemilihan kata, atau kekurangan maka di tahap ini kesalahan dan kekurangan tersebut diperbaiki.

13

4.1.2 Expert Reviews

Pada tahap ini, validasi soal secara kualitatif diperiksa berdasarkan konten, konstruk dan bahasa oleh pembimbing tesis. Selain itu peneliti juga meminta pendapat dari beberapa orang ahli yang sudah berpengalaman sebagai validator ahli.

Adapun validator-validator tersebut adalah :

1. Prof. Dr. Usman Mulbar, M.Pd., selaku dosen Magister Pendidikan Matematika Universitas Negeri Makassar.
2. Dr. Thien Lei Mee, selaku dosen UTM dan R & D Specialist *Mathematics Education SEAMEO RECSAM* Gelugor, Penang, Malaysia.
3. Kery Cotter, selaku *research specialist and TIMSS Mathematic coordinator*.

Karena keterbatasan waktu, proses validasi dilakukan melalui *email* kepada ketiga validator yaitu Prof. Dr. Usman Mulbar, M.Pd., selaku dosen Magister Pendidikan Matematika Universitas Negeri Makassar. Dr. Thien Lei Mee, selaku dosen UTM dan R & D Specialist *Mathematics Education SEAMEO RECSAM* Gellugor, Penang, Malaysia, dan Kery Cotter, selaku *research specialist and TIMSS Mathematic coordinator*.

4.1.3 One to one

Pada tahap ini, soal-soal yang dikembangkan pada *Prototype I* diujikan kepada 3 orang siswa SD Islam Az-Zahrah Palembang dengan tingkat kemampuan yang berbeda. Ketiga siswa tersebut adalah IP dengan kemampuan tinggi, RS dengan kemampuan sedang, dan M dengan kemampuan rendah.

Berdasarkan dari hasil hasil uji validitas kuantitatif, ada lima soal dinyatakan tidak valid. Sedangkan untuk koefisien reliabilitas dari soal-soal *Prototype I*

diperoleh nilai α sebesar 0,74 yang menunjukkan bahwa soal-soal *TIMSS* tipe berpikir tingkat tinggi yang valid dan praktis untuk kelas IV memiliki reliabilitas tinggi karena nilainya nilai $\alpha > 0,60$.

14

4.1.3 Small Group

Soal-soal yang telah direvisi berdasarkan *expert review* dan *one-to-one* dinamakan *prototype 2*. Soal-soal tersebut diujicobakan pada *small group* yang terdiri dari 6 orang siswa SD Islam Az-Zahrah Palembang dengan kemampuan berbeda, yaitu 3 siswa dengan kemampuan tinggi, dua siswa dengan kemampuan sedang dan satu siswa dengan kemampuan rendah.

4.1.4 Field test

Field test dilaksanakan di kelas IV Fatonah SD Islam Az-Zahrah Palembang dengan jumlah siswa 23 orang yang terdiri dari 12 orang siswa laki-laki dan 11 orang siswa perempuan, namun 4 orang siswa yang tidak hadir sehingga penelitian dilakukan dengan 19 orang siswa.

Field test ini bertujuan untuk mengetahui efek potensial soal *TIMSS* tipe kemampuan berpikir tingkat tinggi. Soal-soal yang diberikan pada *field test* ini merupakan 10 soal yang telah valid dan praktis yang telah dikembangkan setelah melalui proses revisi dari hasil *one to one*, *expert review*, dan *small group*. Sebelum pelaksanaan *test*, peneliti memberikan pengarahan mengenai tata cara pelaksanaan *test* serta memfasilitasi siswa dengan seperangkat soal dengan lembar jawaban yang dikumpulkan pada saat tes berakhir

Proses pengembangan soal terdiri dari tahap *self evaluation*, *expert reviews* dan *one to one*, *small group*, dan *field test* serta revisi pada masing-masing ini telah menghasilkan seperangkat soal matematika model *TIMSS* yang valid dan praktis. Seperangkat soal tersebut dinyatakan valid dan praktis secara kualitatif oleh pakar melalui melalui proses validasi dari beberapa validator yang memberikan kontribusi berupa saran dan komentar terhadap perbaikan soal baik dari segi konten, konstruk dan bahasa.

5. KESIMPULAN DAN SARAN

5.1 Kesimpulan

Telah dihasilkan perangkat soal tipe berpikir tingkat tinggi model TIMSS untuk siswa kelas IV SD sebanyak 10 butir soal yang valid dan praktis. Valid terlihat dari validasi pakar dan validasi butir soal. Valid secara kualitatif tergambar dari hasil penilaian validator, dimana semua validator menyatakan baik dan hasil dari *one to one*. Adapun valid secara kuantitatif tergambar berdasarkan analisis butir soal (validitas butir soal) yang bukan subjek penelitian, sedangkan praktis tergambar dari ujicoba terhadap kelompok kecil (*small group*) yang dapat menggunakan perangkat soal dengan baik. Jadi, kesepuluh soal dikategorikan valid dan praktis secara kualitatif maupun secara kuantitatif.

5.2 Saran

Berdasarkan hasil penelitian dan kesimpulan di atas, maka disarankan:

1. Bagi guru dapat menerapkan dalam pembelajaran matematika di sekolah sebagai apresiasi dalam perbaikan evaluasi pembelajaran.
2. Bagi siswa dapat menggunakan soal-soal model TIMSS ini sebagai alternatif karena soal-soal matematika model TIMSS dapat membantu siswa untuk menunjukkan pengetahuan siswa, menerapkan apa yang siswa telah pelajari, menganalisis, mengevaluasi suatu persoalan, dan membantu siswa memiliki kemampuan mengkreasi dalam kehidupan sehari-hari.
3. Bagi peneliti lain diharapkan dapat dipergunakan sebagai bahan masukan untuk mendesain soal-soal model TIMSS diseluruh sekolah dan bisa diaplikasikan dengan materi-materi yang lain dalam pembelajaran matematika.

DAFTAR PUSTAKA

- Buku : De Lange, J. 1999. *Framework For Classroom Assessment in Mathematics*, Freudentagl Institute & National Center for improving Student Learning and Achievement in Mathematics and Science.
- Buku : Djaali & Muljono, P. 2008. *Pengukuran dalam Bidang Pendidikan*.

Jakarta: PT. Grasindo.

- Buku : Krathwohl. 1997. *Methods of Educational and Social Science Research, Second Edition*. New York: Longman, Inc
- Buku : Lie, Thien. 2015. *Higher Order Thinking Skills and International Large Scale Assessment*. Malaysia: UTM.
- Tesis : Lewy. 2009. *Pengembangan Soal Untuk Mengukur Kemampuan Berpikir Tingkat Tinggi Pokok Bahasan Barisan dan Deret Bilangan di Kelas Akselerasi SMP Xaverius Maria Palembang*. Palembang :Universitas Sriwijaya.
- Buku : Mullis, I.V.S., Martin, M.O., Foy, P., & Arora, A. 2012. *TIMSS 2011 Internastional Result in Mathematics*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
- Buku : Tessmer, Martin. 1998. *Planning and conducting formative evaluations: improving the quality of education and training*. London: Kogan Page.
- Buku : Wardhani, S dan Rumiati. 2011. *Modul Matematika SMP Program BERMUTU. Instrumen Penilaian Hasil Belajar Matematika SMP : Belajar dari PISA dan TIMSS*. Yogyakarta:Pusat Pengembangan dan Pemberdayaan Pendidik dan Tenaga Kependidikan (PPPPTK) Matematika.
- Buku : Zulkardi, Z. 2002. *Developing a learning environment on realistic mathematics edu-cation for Indonesian student teachers*. Doctoral dissertation. Enschede: Uni-versity of Twente, Enschede.The Netherland.(Online). Tersedia: http://doc.utwente.nl/58718/1/thesis_Zulkardi.pdf. diakses tanggal 11 November 2015.

