



Skripsi

**PENGEMBANGAN INSTRUMEN ASESMEN *TESTLET*  
BERBASIS *SCIENCE, TECHNOLOGY, ENGINEERING,*  
*AND MATHEMATICS* PADA MATERI TEKANAN ZAT**

diajukan sebagai salah satu syarat untuk memperoleh gelar Sarjana Pendidikan  
Program Studi Pendidikan IPA

Oleh  
Shofie Maulani  
4001415047

**JURUSAN IPA TERPADU  
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM  
UNIVERSITAS NEGERI SEMARANG  
2019**

## PERNYATAAN

Saya menyatakan bahwa skripsi ini bebas plagiat, dan apabila di kemudian hari terbukti terdapat plagiat dalam skripsi ini, maka saya bersedia menerima sanksi sesuai ketentuan peraturan perundang-undangan.

Semarang, Juni 2019



Shofie Maulani

4001415047

## PENGESAHAN

Skripsi yang berjudul

PENGEMBANGAN INSTRUMEN ASESMEN *TESTLET* BERBASIS  
*SCIENCE, TECHNOLOGY, ENGINEERING, AND MATHEMATICS* PADA  
MATERI TEKANAN ZAT

disusun oleh

Shofie Maulani

4001415047

telah dipertanyakan dihadapan sidang Panitia Ujian Skripsi FMIPA UNNES pada  
tanggal 24 Juni 2019.

Panitia:



Ketua  
Prof. Dr. Sudarmin, M.Si.

NIP 1906001231992031003

Sekretaris

Novi Ratna Dewi, S.Pd., M.Pd.

NIP 198311102008012008

Ketua Penguji

Penguji Utama

Stephani Diah Pamelasari, S.S., M.Hum.

NIP 198505142010122007

Anggota Penguji I

Penguji Pendamping

Miranita Khusniati, M.Pd.

NIP 198511162012122003

Anggota Penguji/

Dosen Pembimbing

Prof. Dr. Wiyanto, M.Si.

NIP 196310121988031001

## **MOTTO DAN PERSEMBAHAN**

### **MOTTO**

“Tuliskan rencana kita dengan sebuah pensil, tetapi berikan penghapusnya kepada Allah. Izinkan Dia menghapus bagian-bagian yang salah dan menggantinya dengan rencana-Nya yang indah di dalam hidup kita, karena Allah selalu tahu apa yang kita butuhkan, bukan apa yang kita minta, dan Allah tidak henti-hentinya memenuhi kebutuhan seseorang, selama ia memenuhi kebutuhan saudaranya.”

**(HR. Thabrani)**

### **PERSEMBAHAN**

Skripsi ini Saya persembahkan untuk:

1. My beloved Ummi, Ibu, Bapak, Adik, dan seluruh keluarga besar Bani Mursidi yang senantiasa memberikan doa, memotivasi, dan mendukung untuk menyelesaikan pendidikan di Universitas Negeri Semarang.
2. Almamater jurusan IPA Terpadu di Universitas Negeri Semarang.

## PRAKATA

Puji syukur hanya kepada Allah SWT, *Rabb* semesta alam yang telah menganugerahkan akal dan hati kepada manusia sehingga penulis dapat menyelesaikan penulisan skripsi ini. Penulisan skripsi melibatkan beberapa pihak yang sangat membantu sehingga penulis mengucapkan terima kasih kepada yang terhormat:

1. Prof. Dr. Sudarmin, M.Si. selaku Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Negeri Semarang.
2. Novi Ratna Dewi, S.Pd., M.Pd. selaku Kepala Jurusan IPA Terpadu.
3. Prof. Dr. Wiyanto, M.Si. selaku dosen pembimbing yang sangat membantu dalam membimbing penulisan skripsi.
4. Fidia Fibriana, M.Si., M.Sc. selaku dosen wali yang selalu mengingatkan dan menyemangati untuk terus berprestasi.
5. Bapak dan Ibu dosen Jurusan IPA Terpadu yang telah mendidik dan membimbing dengan ilmu yang bermanfaat bagi penulis selama menjadi mahasiswa di Universitas Negeri Semarang.
6. Prof. Dr. Ani Rusilowati, M.Pd., Dr. Endang Susilaningih, M.S., Muhamad Taufiq, M.Pd., Prasetyo Listiaji, M.Pd., Fidia Fibriana, M.Sc., dan Erna Noor Savitri, S.Si., M.Pd., selaku validator soal *testlet* yang dikembangkan penulis.
7. Suhirman, M.Pd. selaku Kepala SMP N 8 Salatiga dan Suyudi, M.Pd. selaku Kepala SMP N 3 Salatiga yang telah mengizinkan penulis untuk menyelesaikan penelitian.
8. Drs. Abadi selaku guru IPA di SMP N 8 Salatiga dan Chatarina Titik Setiyarini, S.Si. selaku SMP N 3 Salatiga yang telah membantu penulis untuk menggunakan jam mengajarnya dalam menyelesaikan penelitian.
9. Dina, Kiki, Widhi, Widiya, Mae, dan teman-teman Kos Asri yang selalu memberi dukungan semangat serta motivasi selama menyusun skripsi.

10. Ni'matullah Handayani, Dian Emy Mastura, Nila Prafitia Sari, dan Septia Nurkhalisa selaku teman seperjuangan yang bersama-sama saling memotivasi dan menguatkan selama kuliah dan menyusun skripsi.
11. Teman-teman Pendidikan IPA angkatan 2015, terima kasih atas segala macam motivasi, semangat, dan kerja sama untuk menyelesaikan skripsi.

Terima kasih atas doa, motivasi, dan dukungan dari semua pihak semoga kebaikan kalian dibalas oleh Allah SWT dengan cara terbaik. Penulis menyadari bahwa skripsi ini tidak sempurna karena terbatasnya kemampuan dan pengetahuan yang penulis miliki, sehingga penulis membutuhkan kritik dan saran yang bersifat membangun dari pembaca sebagai bahan evaluasi. Kekurangan dan kelemahan yang ada pada penulisan skripsi menjadi harapan bagi penulis sebagai sumber inspirasi pembaca untuk menulis dengan cara yang lebih baik lagi sehingga dapat memberi manfaat dan menambah pengetahuan bagi semuanya.

Semarang, Juni 2019

Penulis

Shofie Maulani

NIM 4001415047

## ABSTRAK

Maulani, S. 2019. *Pengembangan Instrumen Asesmen Testlet Berbasis Science, Technology, Engineering, and Mathematics pada Materi Tekanan Zat*. Skripsi, Jurusan IPA Terpadu Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Negeri Semarang. Pembimbing Utama Prof. Dr. Wiyanto, M.Si.

**Kata kunci:** pengembangan, *testlet*, STEM, tekanan zat

Instrumen asesmen digunakan sebagai alat ukur hasil belajar peserta didik setelah mengikuti pembelajaran. Pembelajaran berpendekatan STEM (*Science, Technology, Engineering, and Mathematics*) belum memiliki standar penilaian baku yang ditetapkan. Pembelajaran STEM membutuhkan pengembangan instrumen asesmen yang terintegrasi disiplin ilmu STEM secara menyeluruh (Harwell *et al.*, 2012:66). SMP Negeri 8 Salatiga telah menerapkan pembelajaran STEM dan membutuhkan instrumen asesmen. *Testlet* dapat dikembangkan sebagai instrumen asesmen sesuai kebutuhan di sekolah tersebut. Penelitian ini bertujuan untuk mengembangkan instrumen asesmen *testlet* berbasis *Science, Technology, Engineering, and Mathematics* pada materi tekanan zat yang layak. Metode penelitian yang digunakan yaitu metode *Research dan Development* (R&D) berdasarkan Sugiyono (2017). *Testlet* berbasis STEM memenuhi karakteristik alat evaluasi yang baik menurut Arifin (2012:65). Hasil penelitian menunjukkan bahwa *testlet* yang dikembangkan layak dan valid digunakan sebagai instrumen asesmen berbasis STEM berdasarkan penilaian oleh validator ahli. Reliabilitas soal *testlet* berbasis STEM adalah 0,866 dengan kategori sangat reliabel. Respon peserta didik dan guru terhadap instrumen asesmen berbasis STEM yang dikembangkan adalah positif. *Testlet* yang telah dikembangkan dilengkapi dengan analisis untuk mengukur daya serap peserta didik sehingga dapat digunakan guru sebagai alternatif asesmen pendeteksi kesulitan belajar peserta didik.

## ABSTRACT

Maulani, S. 2019. *Pengembangan Instrumen Asesmen Testlet Berbasis Science, Technology, Engineering, and Mathematics pada Materi Tekanan Zat*. Skripsi, Jurusan IPA Terpadu Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Negeri Semarang. Pembimbing Utama Prof. Dr. Wiyanto, M.Si.

**Keywords:** development, testlet, STEM, pressure substances

Assessment instruments are used as a measurement of students' learning outcomes. Learning with STEM (Science, Technology, Engineering, and Mathematics) approach does not yet have a standardized standard of assessment. STEM learning requires the development of assessment instruments that integrate STEM disciplines as a whole (Harwell et al., 2012: 66). SMP Negeri 8 Salatiga has implemented STEM learning and requires assessment instruments. Testlet can be developed as an assessment instrument that necessary for the school. The research aims to develop assessment instruments testlet-based Science, Technology, Engineering, and Mathematics on decent substance pressure materials. The research method used is the method of Research and Development (R&D) based on Sugiyono (2017). Assessment instruments Testlet-based Science, Technology, Engineering, and Mathematics are valid to use. The testlet based STEM that have been developed have a characteristic evaluation tool according to Arifin (2012:65). The results of the study showed that the testlets that are developed is valid to used as an assessment STEM based on the assesing by expert validators. The reliability value of testlet-based STEM is 0.866 with a highly reliable category. Students' and teacher responses to the development of assessment instruments testlet-based STEM are positive. The testlet that has been developed is equipped with analyses to measure the absorption of students, so testlet based-STEM can be used by teachers as an alternative assessment of students' diagnostic test.



## DAFTAR ISI

Halaman Judul.....	i
PERNYATAAN.....	ii
PENGESAHAN .....	iii
MOTTO DAN PERSEMBAHAN .....	iv
PRAKATA.....	v
ABSTRAK .....	vii
ABSTRACT.....	viii
DAFTAR ISI.....	ix
DAFTAR TABEL.....	xii
DAFTAR GAMBAR .....	xiii
DAFTAR LAMPIRAN.....	xiv
BAB 1 PENDAHULUAN .....	1
1.1 Latar Belakang .....	1
1.2 Rumusan Masalah .....	4
1.3 Tujuan Penelitian.....	4
1.4 Manfaat Penelitian.....	5
1.4.1 Manfaat Teoretis .....	5
1.4.2 Manfaat Praktis .....	5
1.4.2.1 <i>Bagi Peserta Didik</i> .....	5
1.4.2.2 <i>Bagi Guru</i> .....	5
1.4.2.3 <i>Bagi Sekolah</i> .....	5
1.4.2.4 <i>Bagi Peneliti</i> .....	5
1.5 Penegasan Istilah .....	6
1.5.1 Pengembangan .....	6
1.5.2 Instrumen Asesmen.....	6
1.5.3 STEM .....	7
1.5.4 Asesmen <i>Testlet</i> Berbasis STEM.....	7
1.5.5 Tekanan Zat.....	7
BAB 2 TINJAUAN PUSTAKA .....	10
2.1 Kajian Pustaka.....	10

2.1.1	<i>Testlet</i> .....	10
2.1.2	STEM .....	13
2.1.3	Instrumen Asesmen .....	15
2.1.4	Tinjauan Materi Tekanan Zat .....	21
2.2	Kajian Penelitian yang Relevan .....	23
2.3	Kerangka Berpikir .....	25
BAB 3 METODE PENELITIAN.....		28
3.1	Lokasi dan Waktu Penelitian.....	28
3.2	Desain Penelitian .....	28
3.2.1	Tahap Pendahuluan ( <i>Define</i> ).....	30
3.2.2	Tahap Pengembangan ( <i>Develop</i> ) .....	31
3.2.3	Tahap Evaluasi ( <i>Evaluate</i> ) .....	33
3.3	Sampel Penelitian .....	33
3.4	Metode Pengumpulan Data .....	33
3.4.1	Studi Pendahuluan.....	33
3.4.2	Metode Validasi Ahli .....	34
3.4.3	Metode Tes.....	34
3.5	Instrumen Penelitian .....	34
3.5.1	Lembar Angket Validasi Ahli .....	35
3.5.2	Lembar Angket Tanggapan Guru.....	35
3.5.3	Lembar Angket Respon Peserta Didik .....	35
3.6	Metode Analisis Data .....	35
3.6.1	Analisis Angket Validasi Ahli .....	35
3.6.2	Analisis Validitas Butir Soal .....	36
3.6.3	Analisis Reliabilitas Butir Soal .....	37
3.6.4	Analisis Daya Beda Butir Soal.....	38
3.6.5	Analisis Indeks Kesukaran Butir Soal.....	39
3.6.6	Analisis Angket Tanggapan Guru .....	39
3.6.7	Analisis Angket Respon Peserta Didik .....	40
3.7	Kriteria Keberhasilan Pengembangan .....	41
BAB 4 HASIL PENELITIAN DAN PEMBAHASAN .....		42
4.1	Hasil Penelitian.....	42

4.1.1	Hasil Kesahihan Instrumen Asesmen.....	42
4.1.2	Respon Peserta Didik dan Guru terhadap Instrumen Asesmen .....	47
4.2	Pembahasan .....	49
4.2.1	Tahap Pendahuluan ( <i>Define</i> ).....	50
4.2.2	Tahap Pengembangan ( <i>Develop</i> ) .....	51
4.2.3	Tahap Evaluasi ( <i>Evaluate</i> ).....	59
4.2.4	Kesahihan Instrumen Asesmen <i>Testlet</i> Berbasis STEM.....	60
4.3	Respon Guru dan Peserta Didik .....	68
4.3.1	Respon guru .....	68
4.3.2	Respon peserta didik .....	69
BAB 5 SIMPULAN DAN SARAN.....		70
5.1	Simpulan.....	70
5.2	Saran.....	70
DAFTAR PUSTAKA .....		72
LAMPIRAN		

## DAFTAR TABEL

Tabel 2.1 Definisi <i>Testlet</i> .....	10
Tabel 2.2 Rubrik Penilaian Jawaban Modifikasi Model GRM.....	12
Tabel 2.3 Definisi STEM .....	13
Tabel 2.4 Dimensi Proses Berpikir .....	20
Tabel 3.1 Kriteria Penilaian Angket Validasi .....	36
Tabel 3.2 Interpretasi terhadap Reliabilitas .....	38
Tabel 3.3 Klasifikasi Daya Pembeda Soal .....	39
Tabel 3.4 Kriteria Indeks Kesukaran Soal .....	39
Tabel 3.5 Kriteria Penilaian Angket Tanggapan Guru.....	40
Tabel 3.6 Kriteria Penilaian Angket Respon Peserta Didik.....	40
Tabel 4.1 Rekapitulasi Kelayakan Asesmen <i>Testlet</i> Berbasis STEM oleh Ahli... 42	
Tabel 4.2 Rekomendasi Ahli.....	44
Tabel 4.3 Hasil Uji Validitas Butir .....	45
Tabel 4.4 Hasil Uji Reliabilitas <i>Testlet</i> Berbasis STEM.....	45
Tabel 4.5 Rekapitulasi Daya Beda <i>Testlet</i> pada Uji Skala Terbatas .....	46
Tabel 4.6 Rekapitulasi Daya Beda <i>Testlet</i> pada Uji Skala Lebih Luas.....	46
Tabel 4.7 Rekapitulasi Indeks Kesukaran <i>Testlet</i> pada Uji Skala Terbatas.....	47
Tabel 4.8 Rekapitulasi Hasil Angket Respon Peserta Didik.....	48
Tabel 4.9 Rekapitulasi Hasil Angket Respon Guru .....	48
Tabel 4.10 Soal <i>Testlet</i> berbasis STEM .....	64
Tabel 4.11 Rekapitulasi Klarifikasi Jawaban <i>Testlet</i> .....	67

## DAFTAR GAMBAR

Gambar 2.1 Contoh Model <i>Testlet</i> Shiell & Slepko (2015:204).....	11
Gambar 2.2 Pola integrasi STEM .....	14
Gambar 2.3 Bagan Alur Kerangka Berpikir .....	27
Gambar 3.1 Langkah-langkah Penelitian dan Pengembangan.....	29
Gambar 4.1 Konsistensi Validasi Instrumen Asesmen <i>Testlet</i> oleh Ahli .....	43
Gambar 4.2 Gambar balok sebelum dan sesudah direvisi .....	53
Gambar 4.3 Panjang kalimat pilihan ganda sebelum dan sesudah direvisi.....	54
Gambar 4.4 Spesifikasi pilihan ganda sebelum dan sesudah direvisi .....	54
Gambar 4.5 Simbol dan tanda operasional sebelum dan sesudah direvisi.....	55
Gambar 4.6 Pilihan ganda disusun paralel sebelum dan sesudah direvisi .....	56
Gambar 4.7 Petunjuk pengerjaan sebelum dan sesudah direvisi .....	57
Gambar 4.8 Halaman judul sebelum dan sesudah direvisi .....	57
Gambar 4.9 Pilihan Ganda no 1.2 sebelum dan sesudah direvisi .....	59
Gambar 4.10 Konsistensi Reliabilitas <i>Testlet</i> Berbasis STEM.....	61
Gambar 4.11 Hasil Pengembangan <i>Testlet</i> .....	65

## DAFTAR LAMPIRAN

Lampiran 1 Silabus Mata Pelajaran Ilmu Pengetahuan Alam (IPA) .....	76
Lampiran 2 Angket Validasi Ahli Asesmen .....	82
Lampiran 3 Angket Validasi Ahli Materi .....	91
Lampiran 4 Angket Validasi Ahli Bahasa .....	100
Lampiran 5 Rekapitulasi Hasil Kelayakan <i>Testlet</i> Berbasis STEM.....	109
Lampiran 6 Analisis Validitas, Reliabilitas, DP, dan IK Uji Skala Terbatas .....	114
Lampiran 7 Analisis Validitas, Reliabilitas, DP, dan IK Uji Skala Lebih Luas .	117
Lampiran 8 Analisis Validitas, Reliabilitas, DP, dan IK Uji Pemakaian.....	121
Lampiran 9 Angket Respon Peserta Didik.....	124
Lampiran 10 Angket Respon Guru .....	126
Lampiran 11 Rekapitulasi Angket Respon Peserta Didik.....	129
Lampiran 12 Lembar Koreksi Jawaban .....	131
Lampiran 13 Lembar Jawab Peserta Didik .....	132
Lampiran 14 Instrumen Asesmen <i>Testlet</i> Berbasis STEM .....	134
Surat Keputusan Dosen Pembimbing.....	165
Surat Izin Penelitian .....	166
Surat Keterangan Penelitian dari Sekolah.....	167
Dokumentasi Penelitian .....	168

# **BAB 1 PENDAHULUAN**

## **1.1 Latar Belakang**

Permendikbud Nomor 23 Tahun 2016 tentang Standar Penilaian Pendidikan, menjelaskan bahwa penilaian hasil belajar peserta didik mencakup kompetensi sikap, pengetahuan, dan keterampilan yang dilakukan secara berimbang. Hal ini menunjukkan bahwa implementasi Permendikbud Nomor 66 Tahun 2013 tentang Standar Penilaian Pendidikan membawa dampak terhadap proses penilaian, yang di dalamnya termasuk model dan teknik serta prosedur penilaian dalam pembelajaran. Fungsi penilaian juga didasarkan pada Undang-Undang Nomor 20 Tahun 2003 pasal 58 ayat (1) adalah untuk memantau proses, kemajuan dan perbaikan hasil belajar peserta didik secara berkesinambungan. Penilaian hasil belajar dilakukan oleh pendidik berdasarkan panduan dari satuan pendidikan dan pemerintah untuk dipertanggung jawabkan kredibilitasnya. Pendidik melakukan penilaian menggunakan alat yang digunakan untuk mengumpulkan informasi tentang hasil belajar peserta didik yang disebut instrumen asesmen.

Instrumen menurut Trianto (2013:129) merupakan alat yang dapat digunakan untuk mengukur tingkat pencapaian kompetensi. Nurhadi (2009:62) menjelaskan pengertian *assessment* adalah proses pengumpulan berbagai data yang bisa memberikan gambaran perkembangan belajar peserta didik. Oleh karena itu, instrumen asesmen adalah alat evaluasi dalam proses pengumpulan berbagai data untuk mengukur tingkat pencapaian kompetensi yang bisa memberikan gambaran perkembangan belajar peserta didik. Susetyo (2015:36) menyatakan penggunaan *assessment* (asesmen) dilaksanakan pada bagian terakhir pembelajaran untuk mengetahui ketercapaian tujuan pembelajaran. Instrumen asesmen digunakan untuk menyelidiki pemahaman peserta didik tentang konsep pengetahuan dan sebagai sarana membuat hubungan antar konsep pengetahuan tersebut (Asmalia *et al.*, 2015:300). Hal ini menunjukkan asesmen tentang pemahaman konseptual sangat penting digunakan dalam proses penilaian pembelajaran.

Penilaian pembelajaran dilakukan ketika proses kegiatan belajar mengajar maupun di akhir pembelajaran. Begitu pula dengan penilaian pada pembelajaran IPA di SMP/MTs. Pembelajaran IPA di SMP/MTs mencakup empat bidang kajian, meliputi energi dan perubahannya, makhluk hidup dan proses kehidupan, materi dan sifatnya, serta bumi antariksa (Ismayani, 2016:268). Penyampaian materi IPA selaras dengan pembelajaran yang sudah berjalan yaitu pembelajaran terpadu. Pembelajaran IPA dengan konsep terpadu telah banyak diimplementasikan bersama dengan pendekatan, metode, dan model yang menjadi *trand* di dunia pendidikan. Salah satu implementasi pembelajaran IPA yang sedang *trend* saat ini yaitu pembelajaran IPA berpendekatan *Science, Technology, Engineering, and Mathematics* yang selanjutnya disingkat STEM. Becker & Park (2011:24) menyatakan pembelajaran terintegrasi STEM dalam hasil analisisnya akan memberikan efektivitas tinggi terhadap pembelajaran peserta didik. Afriana *et al.*, (2016:203) telah mengimplementasikan pembelajaran IPA dengan model *Project Based Learning* terintegrasi STEM.

Fakta tentang implementasi pembelajaran IPA berbasis STEM diperkuat berdasarkan hasil observasi di SMP Negeri 8 Salatiga. Implementasi pembelajaran berpendekatan STEM di SMP Negeri 8 Salatiga dilakukan sejak 2 tahun terakhir. Hasil observasi di SMP Negeri 8 Salatiga menemui beberapa permasalahan. Masalah yang muncul diantaranya yaitu, kemampuan akademik lebih rendah dibanding kemampuan non-akademik, rendahnya keterampilan berpikir tingkat tinggi peserta didik, instrumen asesmen yang digunakan masih menggunakan instrumen lama dan penurunan hasil belajar pada materi IPA terkait hitungan. Penurunan nilai hasil belajar peserta didik kelas VIII paling signifikan pada materi hitungan, contohnya materi semester I tentang gaya dan usaha; materi semester II tentang materi tekanan zat. Rahmawati *et al.*, (2016:102) mengungkapkan materi tekanan zat dan penerapannya dalam kehidupan merupakan materi abstrak yang sulit dipahami konsepnya.

Identifikasi lebih lanjut tentang faktor yang mempengaruhi penurunan hasil belajar peserta didik serta sistem penilaian pembelajaran STEM yang



digunakan. Penilaian asesmen autentik dari guru menunjukkan nilai keterampilan dan psikomotorik peserta didik sangat baik, penurunan terbesar ada pada hasil penilaian aspek kognitif. Penilaian aspek kognitif yang digunakan guru adalah tes berupa soal 10 soal pilihan ganda dan 4 soal uraian yang belum berbasis STEM. Instrumen asesmen yang digunakan untuk mengukur kemampuan hasil belajar harus sesuai dengan proses pembelajaran yang diterima peserta didik.

Pendekatan STEM di Indonesia masih dalam tahap berkembang, dan saat ini belum terdapat instrumen penilaian baku yang dapat diterapkan khusus untuk pembelajaran STEM. Malik *et al.*, (2018:12) mengembangkan instrumen asesmen HOTS (*Higher Order Thinking Skills*) menggunakan model inkuiri terbimbing sebagai alat evaluasi pembelajaran berbasis inkuiri. Begitu pula dengan pembelajaran berpendekatan STEM, penilaian yang digunakan juga menggunakan instrumen asesmen berbasis STEM. Harwell *et al.*, (2012:66) menyebutkan model baru pembelajaran STEM membutuhkan pengembangan instrumen asesmen yang terintegrasi disiplin ilmu STEM secara menyeluruh. Fakta ini menunjukkan benar diperlukan pengembangan instrumen asesmen berbasis STEM. Hasil pengumpulan data melalui observasi dan wawancara di SMP Negeri 8 Salatiga diperlukan pengembangan instrumen penilaian pembelajaran berbasis STEM. Menanggapi pernyataan tersebut, peneliti bermaksud pengembangan instrumen asesmen berbasis STEM dalam bentuk tes objektif.

Instrumen asesmen bentuk tes objektif telah banyak dikembangkan. Salah satu bentuk pengembangan bentuk tes objektif adalah instrumen asesmen model *testlet*. *Testlet* adalah kumpulan soal terpadu (*Integrated Testlets-ITs*) yang terdiri dari empat item pilihan ganda yang masing-masing mewakili langkah penting dalam menyelesaikan masalah (Shiell & Slepko, 2015:205). Butir soal berjenjang dalam *testlet* dianggap sebagai satu grup penilaian yang berbagi permasalahan pada satu konteks (Scalise & Wilson, 2007:2). Jawaban yang benar pada setiap soal dapat menyampaikan kepada peserta didik secara keseluruhan, sebagian, ataupun tidak sebagai tanda yang didapat sebelum

mereka melanjutkan soal selanjutnya dengan pengetahuan penuh untuk jawaban benar (Shiell & Slepko, 2015:205). Nova *et al.*, (2016:1197) mengembangkan instrumen asesmen dengan menggabungkan efisiensi tes pilihan ganda dan uraian berupa instrumen asesmen *testlet*. *Testlet* merupakan model tes yang sesuai untuk dikembangkan sebagai asesmen berbasis STEM di SMP Negeri 8 Salatiga karena kebutuhan instrumen asesmen yang mempermudah guru mengidentifikasi dan melatih kemampuan berpikir tingkat tinggi peserta didik.

Merujuk pada capaian penelitian yang telah dipaparkan diatas, terdapat celah penelitian berupa kebutuhan pengembangan instrumen asesmen pada pembelajaran STEM. Penelitian ini fokus pada pengembangan instrumen asesmen *testlet* berbasis *Science, Technology, Engineering, and Mathematics* pada materi tekanan zat.

## **1.2 Rumusan Masalah**

Berdasarkan latar belakang, rumusan masalah yang akan dikaji dalam penelitian ini adalah:

1. Bagaimana validitas, reliabilitas, daya beda, dan indeks kesukaran instrumen asesmen *testlet* berbasis *Science, Technology, Engineering, and Mathematics* sebagai lembar asesmen pada materi tekanan zat?
2. Bagaimana respon peserta didik dan guru terhadap instrumen asesmen *testlet* berbasis *Science, Technology, Engineering, and Mathematics* pada materi tekanan zat?

## **1.3 Tujuan Penelitian**

Berdasarkan rumusan masalah diatas, tujuan dari penelitian ini adalah:

1. Menganalisis validitas, reliabilitas, daya beda, dan indeks kesukaran instrumen asesmen *testlet* berbasis *Science, Technology, Engineering, and Mathematics* sebagai lembar asesmen pada materi tekanan zat.
2. Menganalisis respon peserta didik dan guru terhadap instrumen asesmen *testlet* berbasis *Science, Technology, Engineering, and Mathematics* pada materi tekanan zat.

## **1.4 Manfaat Penelitian**

Manfaat yang dapat diambil dalam penelitian ini terdapat dua macam yaitu manfaat teoretis dan praktis:

### **1.4.1 Manfaat Teoretis**

Hasil penelitian pengembangan ini dapat memberikan sumber referensi pada bidang pendidikan berupa artikel ilmiah tentang pengembangan instrumen asesmen *testlet* berbasis STEM pada materi tekanan zat.

### **1.4.2 Manfaat Praktis**

#### **1.4.2.1 Bagi Peserta Didik**

Asesmen *testlet* berbasis STEM materi tekanan zat dapat melatih dan mengukur kemampuan peserta didik melalui latihan mengerjakan soal pilihan ganda dengan model *testlet*.

#### **1.4.2.2 Bagi Guru**

Guru memiliki instrumen asesmen *testlet* berbasis STEM peserta didik pada materi tekanan zat, dan guru dapat mengembangkan asesmen *testlet* pada materi lain; sebagai bahan refleksi diri terhadap asesmen yang telah diterapkan selama mengajar; memberi alternatif penilaian oleh guru untuk melengkapi tes tertulis yang mempermudah guru memahami karakteristik masing-masing peserta didik, mengevaluasi, dan tindak lanjut yang diberikan.

#### **1.4.2.3 Bagi Sekolah**

Sekolah dapat mensosialisasikan penggunaan instrumen *testlet* berbasis STEM materi tekanan zat dan menggandakannya untuk melengkapi instrumen asesmen *testlet*; sekolah mendapatkan gambaran tentang asesmen yang selama ini dilakukan oleh guru di dalam kelas; serta sekolah dapat mengupayakan solusi, pembinaan dan tindak lanjut terhadap guru yang masih kesulitan melaksanakan asesmen pembelajaran.

#### **1.4.2.4 Bagi Peneliti**

Mendapat pengetahuan, pengalaman, dan wawasan dalam melakukan penelitian, menulis artikel ilmiah, serta melatih diri untuk

menerapkan ilmu pengetahuan tentang pengembangan instrumen *testlet* berbasis STEM pada materi tekanan zat.

## **1.5 Penegasan Istilah**

Dalam penelitian ini perlu dijelaskan istilah yang digunakan untuk mengetahui batasan dalam penelitian. Adapun istilah yang perlu dijelaskan adalah sebagai berikut.

### **1.5.1 Pengembangan**

Pengembangan atau dalam bahasa Inggris *Research and Development* adalah metode penelitian yang digunakan untuk menghasilkan produk tertentu, dan menguji keefektifan produk tersebut (Sugiyono, 2017:297). Langkah-langkah penelitian dan pengembangan menggunakan model Sugiyono (2017:298) yang terbagi menjadi sepuluh langkah yaitu, (1) potensi dan masalah, (2) mengumpulkan informasi, (3) desain produk, (4) validasi desain, (5) perbaikan desain, (6) uji coba produk, (7) revisi produk, (8) uji coba pemakaian, (9) revisi produk, dan (10) pembuatan produk masal. Penelitian pengembangan instrumen asesmen *testlet* berbasis STEM pada materi tekanan zat hanya sampai pada langkah ke-9 yang disederhanakan dalam tiga tahapan utama yaitu pendahuluan, pengembangan dan evaluasi.

### **1.5.2 Instrumen Asesmen**

Instrumen menurut Trianto (2013:129) merupakan alat yang dapat digunakan untuk mengukur tingkat pencapaian kompetensi. Nurhadi (2009:62) menjelaskan pengertian *asesment* adalah proses pengumpulan berbagai data yang bisa memberikan gambaran perkembangan belajar peserta didik. Teknik evaluasi digolongkan menjadi dua yakni teknik tes dan nontes. Rudyatmi & Rusilowati (2012:47) menjelaskan bahwa teknik tes merupakan teknik penilaian untuk mengukur ranah kognitif, afektif, dan psikomotorik peserta didik. Instrumen asesmen dalam penelitian ini fokus pada asesmen tes untuk mengukur kemampuan pada aspek kognitif peserta didik menggunakan model *testlet*.

### 1.5.3 STEM

Tsupros *et al.*, (2009:24) mendefinisikan STEM adalah pendekatan interdisipliner pada pembelajaran dimana peserta didik menggunakan sains, teknologi, teknik, dan matematika dalam konteks nyata yang menghubungkan antara sekolah, bidang pekerjaan, dan masyarakat secara terpadu, sehingga mampu mengembangkan literasi peserta didik untuk diimplementasikan dalam kehidupan. Pendapat lain tentang STEM juga dinyatakan oleh Brown *et al.*, (2011:6) bahwa STEM adalah metadisiplin di sekolah dimana guru sains, teknologi, teknik, dan matematika mengajar pendekatan terpadu dan masing-masing disiplin ilmu ini dipadukan sebagai satu kesatuan yang dinamis. Pengembangan instrumen asesmen yang dikembangkan pada penelitian ini disusun berdasarkan pada indikator sains, teknologi, teknik, dan matematika dalam konteks nyata.

### 1.5.4 Asesmen *Testlet* Berbasis STEM

*Testlet* adalah kumpulan soal terpadu (*Integrated Testlets-ITs*) yang terdiri dari empat item pilihan ganda atau lebih yang masing-masing mewakili langkah penting dalam menyelesaikan masalah (Shiell & Slepko, 2015:205). Instrumen asesmen *testlet* yang akan dikembangkan dalam penelitian ini adalah soal *testlet* berupa kumpulan soal terpadu yang terdiri dari empat item pilihan ganda atau lebih, yang disusun berjenjang dari cara paling mudah mengikuti pola berpikir mengerjakan soal uraian. Pengembangan instrumen ini meliputi menyusun instrumen asesmen yang disesuaikan dengan indikator STEM dalam model *testlet*. Penelitian ini membatasi permasalahan pada kebutuhan terhadap instrumen asesmen *testlet* berbasis STEM untuk dikembangkan menjadi instrumen asesmen yang layak. Adapun instrumen asesmen yang dikembangkan dalam penelitian ini adalah alat evaluasi tes objektif berupa *testlet* berbasis STEM.

### 1.5.5 Tekanan Zat

Tekanan zat merupakan materi pelajaran IPA kelas VIII semester 2 dalam kurikulum 2013 revisi 2017. Materi tekanan zat dibahas dalam bab pertama tentang tekanan zat dan penerapannya dalam kehidupan sehari-hari

sesuai dengan KD 3.8 dan 4.8. Rahmawati *et al.*, (2016:102) mengungkapkan materi tekanan zat dan penerapannya dalam kehidupan merupakan materi abstrak yang sulit dipahami konsepnya. Materi ini membahas tekanan zat yang meliputi tekanan zat padat, tekanan zat cair, tekanan zat gas serta hukum-hukum yang berlaku didalamnya. Materi tekanan zat dalam penelitian ini merupakan tema untuk mata pelajaran IPA Terpadu di SMP/MTs dengan pembelajaran berpendekatan STEM dengan memperhatikan kompetensi dasar mata pelajaran IPA Terpadu di SMP/MTs. Fokus materi pada pengembangan instrumen *testlet* ini pada materi tekanan zat.

## BAB 2 TINJAUAN PUSTAKA

### 2.1 Kajian Pustaka

#### 2.1.1 *Testlet*

Asesmen *testlet* telah banyak dikembangkan sebelumnya dan memiliki beberapa definisi operasional seperti disajikan pada Tabel 2.1.

Tabel 2.1 Definisi *Testlet*

Sumber	Definisi
Slepkov & Ralph (2014:10)	<i>Testlet</i> adalah struktur tes baru yang dikembangkan dan dirancang untuk mewakili pertanyaan uraian dimana prosedur atau langkah pengerjaannya sebagai seperangkat pertanyaan pada pilihan ganda.
Shiell & Slepkov (2015:205)	<i>Testlet</i> adalah kumpulan soal terpadu ( <i>Integrated Testlets-ITs</i> ) yang terdiri dari empat item pilihan ganda yang masing-masing mewakili langkah penting dalam menyelesaikan masalah.
Kusumaningrum <i>et al.</i> , (2015:39)	<i>Testlet</i> adalah model pilihan ganda berjenjang yang memadukan keunggulan bentuk tes model pilihan berganda dan uraian objektif.

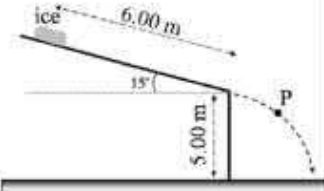
Definisi *testlet* diperjelas oleh Yamtinah *et al.*, (2014:5) sebagai sebuah butir soal yang terdiri dari beberapa soal pendukung yang bersifat dependen. Soal pendukung no 1 menjadi dasar bagi soal-soal pendukung berikutnya sehingga jika soal pendukung no 1 dijawab salah oleh peserta didik, maka peserta didik tidak akan bisa menjawab benar soal-soal pendukung berikutnya. Butir soal berjenjang dalam *testlet* dianggap sebagai satu grup penilaian yang berbagi permasalahan pada satu konteks (Scalise & Wilson, 2007:2). Jawaban yang benar pada setiap soal dapat menyampaikan kepada peserta didik secara keseluruhan, sebagian, ataupun tidak sebagai tanda yang didapat sebelum mereka melanjutkan soal selanjutnya dengan pengetahuan penuh untuk jawaban benar (Shiell & Slepkov, 2015:205).

Berdasarkan definisi-definisi tersebut, instrumen asesmen *testlet* yang akan dikembangkan dalam penelitian ini adalah soal *testlet* berupa kumpulan soal terpadu yang terdiri dari empat item pilihan ganda atau

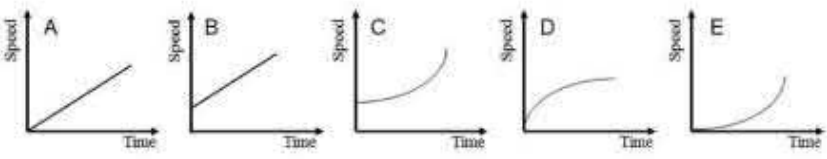
lebih, yang disusun berjenjang dari cara paling mudah mengikuti pola berpikir mengerjakan soal uraian. Model penyusunan *testlet* yang dikembangkan dalam penelitian ini seperti ditunjukkan pada Gambar 2.1.

**Kinematics and projectile motion**

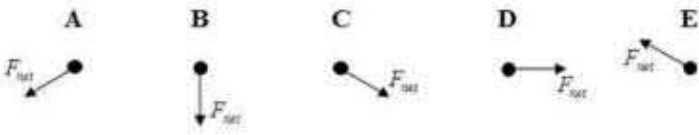
A piece of ice is 6.00 m from the edge of a roof of height 5.00 m at an angle of  $15^\circ$  to the horizontal as shown in the figure, when it begins to slide. The coefficient of kinetic friction between the ice and the roof is 0.2. Ignore any air resistance.



1) Which of the following graphs best represents the speed of the ice as a function of time, beginning from the moment when it first begins to slide and ending when it leaves the roof?



2) Which of the following free-body diagrams is most correct for the ice at position P?



3) How much time elapses between the ice beginning to slide and it leaving the roof?

A. 0.65 s      B. 4.3 s      C. 1.1 s      D. 2.2 s      E. 1.6 s

4) At what (horizontal) distance from the base of the house does the ice land?

A. 5.3 m      B. 1.9 m      C. 19 m      D. 3.5 m      E. 2.5 m

Ans: A:B:B:E

Gambar 2.1 Contoh Model *Testlet* Shiell & Slepko (2015:204)

Soal *testlet* yang ditunjukkan pada Gambar 2.1 menunjukkan bahwa *testlet* tersebut memiliki 4 item soal pendukung yang mewakili langkah penting dalam menyelesaikan soal. Independensi setiap item dalam *testlet* menunjukkan bahwa antar soal pendukung saling terkait. Item 1 menunjukkan pertanyaan pilihan ganda yang berbasis identifikasi grafik hubungan kecepatan-waktu, item 2 peserta didik mengidentifikasi sumber gaya, item 3 peserta didik mengaplikasikan persamaan rumus yang diidentifikasi pada item 1, dan item 4 dapat diselesaikan dengan



menyelesaikan item 1, 2, dan 3 dengan benar (Shiell & Slepkov, 2015:205).

Pedoman penskoran *testlet* menggunakan *Graded Response Model* (GRM) karena *testlet* termasuk soal yang disusun dengan model berjenjang. Yamtinah *et al.*, (2014:5) melakukan penskoran dengan menggunakan GRM yang menerapkan sistem grading untuk menilai soal *testlet*. Pedoman penskoran *testlet* menggunakan rubrik penilaian jawaban model GRM sesuai tabel 2.2 yang diadaptasi dari Murti *et al.*, (2018:27).

Tabel 2.2 Rubrik Penilaian Jawaban Modifikasi Model GRM

No	Aspek Penilaian	Skor
1	Peserta didik tidak dapat menyelesaikan langkah atau tahapan pertama dengan benar	0
2	Peserta didik dapat menyelesaikan langkah atau tahapan pertama dengan benar tetapi tidak dapat menyelesaikan tahap kedua	1
3	Peserta didik dapat menyelesaikan langkah atau tahapan pertama dan kedua dengan benar tetapi tidak dapat menyelesaikan tahap ketiga	2
4	Peserta didik dapat menyelesaikan langkah atau tahapan 1, 2, dan 3 dengan benar	3
5	Peserta didik dapat menyelesaikan langkah atau tahapan kedua ataupun ketiga, tetapi tidak mampu menyelesaikan langkah pertama	skor sesuai dengan jumlah soal yang dijawab benar

Penggunaan *testlet* sebagai instrumen asesmen memiliki keunggulan mampu mengukur kemampuan peserta didik dalam menguasai materi pelajaran. Berdasarkan kajian teoretik dan empirik yang telah dilakukan Susongko (2010:271) ternyata dari segi penskoran, *testlet* lebih praktis dibanding bentuk uraian karena penskoran dapat dilakukan secara objektif dan bersifat politomus. Keunggulan penggunaan tes model *testlet* didukung oleh Yamtinah *et al.*, (2014:2) menyatakan *testlet* yang dilengkapi dengan analisis profil individu peserta didik akan membantu guru sehingga lebih mudah untuk mencermati letak kelemahan peserta didik. Instrumen asesmen *testlet* juga dapat dirancang sebagai suatu instrumen yang dapat mengukur kemampuan kognitif peserta didik karena

dikembangkan berdasarkan subjek spesifik (mata pelajaran) tertentu. Berdasarkan uraian tersebut diketahui bahwa *testlet* dapat digunakan sebagai instrumen asesmen tes formatif yang dapat mengukur kemampuan hasil belajar peserta didik.

### 2.1.2 STEM

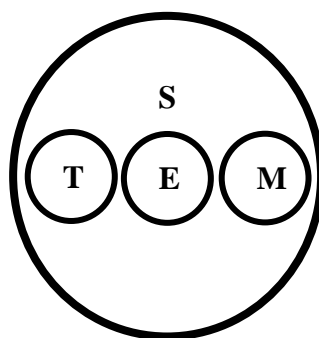
STEM adalah kegiatan belajar dan pembelajaran dalam bidang *science, technology, engineering, and mathematics* (Gonzalez & Kuenzi, 2012:2). Pendekatan STEM merupakan sebuah pendekatan pembelajaran yang menggunakan pendekatan antar ilmu dimana pengaplikasiannya dilakukan dengan pembelajaran aktif berbasis permasalahan (Afriana *et al.*, 2016:203). Definisi pendekatan STEM menurut Ismayani (2016:268) dijelaskan pada Tabel 2.3.

Tabel 2.3 Definisi STEM

Disiplin Ilmu	Definisi
<i>Science</i>	Sains: kemampuan dalam menggunakan kemampuan ilmiah dan proses dalam memahami alam semesta serta kemampuan untuk
<i>Technology</i>	Teknologi: pengetahuan bagaimana menggunakan teknologi baru, memahami bagaimana teknologi baru dikembangkan, dan memiliki kemampuan untuk menganalisis bagaimana teknologi baru mempengaruhi individu masyarakat, bangsa dan
<i>Engineering</i>	Desain atau teknik: pemahaman tentang bagaimana teknologi dapat dikembangkan melalui proses rekayasa/desain menggunakan tema pelajaran berbasis proyek dengan cara mengintegrasikan beberapa mata pelajaran
<i>Mathematics</i>	Matematika: kumpulan dalam menganalisis, alasan, dan mengkomunikasikan ide secara efektif dan dari cara bersikap, merumuska, memecahkan, dan menafsirkan solusi untuk masalah matematika dalam menerapkan berbagai situasi berbeda.

Pola keterpaduan STEM diajarkan kepada peserta didik mata pelajaran terintegrasi dengan cara melebur keempat disiplin

misalnya konten teknologi, desain atau teknik, dan matematika sains. Terdapat beragam cara integrasi yang digunakan dalam mengintegrasikan disiplin-disiplin STEM (Roberts, 2012:3). keempat disiplin ilmu tersebut diperjelas pada Gambar 2.2.



Gambar 2.2 Pola integrasi STEM

Tujuan pembelajaran berpendekatan STEM menurut (Bybee, 2013:5) adalah untuk mengembangkan kemampuan peserta didik yang mencakup:

- a. Pengetahuan, sikap, dan keterampilan untuk mengidentifikasi pertanyaan dan masalah dalam situasi kehidupannya, menjelaskan fenomena alam, mendesain serta menarik kesimpulan berdasarkan bukti mengenai isu-isu terkait STEM.
- b. Memahami karakteristik fitur-fitur disiplin STEM sebagai bentuk-bentuk pengetahuan, penyelidikan, serta desain yang digagas
- c. ~~Kesediaan~~ Mengetahui bagaimana disiplin-disiplin STEM membentuk lingkungan, material, intelektual, dan kultural.
- d. Mau terlibat dalam kajian isu-isu terkait STEM (misalnya efisiensi energi, kualitas lingkungan, keterbatasan sumber daya alam) sebagai warga negara yang konstruktif, peduli, serta reflektif dengan menggunakan gagasan-gagasan sains, teknologi, teknik, dan matematika.

Pembelajaran berpendekatan STEM perlu menekankan beberapa aspek dalam proses pembelajaran (National Research Council, 2012:3-5) diantaranya: (1) mengajukan pertanyaan (*science*) dan mendefinisikan

masalah (*engineering*); (2) mengembangkan dan menggunakan model; (3) merencanakan dan melakukan investigasi; (4) menganalisis dan menafsirkan data (*mathematics*); (5) menggunakan matematika; teknologi informasi dan komputer; dan berpikir komputasi; (6) membangun eksplanasi (*science*) dan merancang solusi (*engineering*); (7) terlibat dalam argumen berdasarkan bukti; (8) memperoleh, mengevaluasi, dan mengkomunikasikan informasi.

Pembelajaran berpendekatan STEM memiliki definisi, pola pendekatan, tujuan dan aspek didalamnya, selain itu asesmen pembelajaran berpendekatan STEM juga dibutuhkan. National Research Council (2012:20) menyatakan bahwa untuk memperbaiki dasar pemikiran K-12 dengan pendekatan STEM, prestasi hasil belajar peserta didik harus diukur dengan akurat dan sekolah harus bertanggung jawab menyelenggarakan pengukuran hasil belajar. Hal ini diperjelas Harwell *et al.*, (2012:66) bahwa model baru pembelajaran STEM membutuhkan pengembangan instrumen asesmen yang terintegrasi disiplin ilmu STEM secara menyeluruh.

### **2.1.3 Instrumen Asesmen**

Instrumen menurut Trianto (2013:129) merupakan alat yang dapat digunakan untuk mengukur tingkat pencapaian kompetensi. Nurhadi (2009:62) menjelaskan pengertian *assessment* adalah proses pengumpulan berbagai data yang bisa memberikan gambaran perkembangan belajar peserta didik. Oleh karena itu, instrumen asesmen adalah alat evaluasi dalam proses pengumpulan berbagai data untuk mengukur tingkat pencapaian kompetensi yang bisa memberikan gambaran perkembangan belajar peserta didik. Gambaran perkembangan belajar peserta didik perlu diketahui oleh guru agar bisa memastikan bahwa peserta didik mengalami proses pembelajaran dengan benar. *Assessment* menekankan proses pembelajaran, maka data yang dikumpulkan harus diperoleh dari kegiatan nyata yang dikerjakan peserta didik pada saat melakukan proses pembelajaran.

Instrumen penilaian yang digunakan untuk mengukur keberhasilan peserta didik menurut Arikunto (2013:47) dibagi menjadi tiga yaitu, tes diagnostik, tes formatif, dan tes sumatif. Instrumen penilaian dinilai penting dalam sebuah pembelajaran karena merupakan sebuah proses pengumpulan data untuk menentukan sejauh mana, dalam hal apa, dan bagian mana tujuan pendidikan sudah tercapai. Hal ini sepemikiran dengan Yusuf (2015:108) dimana guru perlu melakukan pengembangan instrumen evaluasi karena tes objektif hanya mengukur kemampuan kognitif saja, sedangkan keberhasilan proses pembelajaran tidak hanya dilihat dari aspek kognitif tetapi dari tiga aspek penilaian yaitu, aspek kognitif, afektif, dan psikomotor. Selain itu, instrumen asesmen yang selama ini banyak digunakan masih berupa soal yang konvergen pada salah satu aspek saja. Rahmawati *et al.*, (2015:843) menyatakan bahwa penggunaan asesmen yang lebih komprehensif akan memudahkan peserta didik dan guru dalam mencapai tujuan belajar yang ditetapkan. Effiong *et al.*, (2014: 2) mengkategorikan jenis tes apabila ditinjau dari bentuk soalnya menjadi dua, yaitu tes hasil belajar bentuk uraian yang sering dikenal dengan tes subjektif dan tes hasil belajar bentuk objektif.

Tes objektif adalah tes pilihan ganda yang jawabannya harus dipilih dari beberapa kemungkinan jawaban yang telah disediakan (Supranata, 2012:107). Stankous (2016:311) menyatakan kelebihan penggunaan tes pilihan ganda dapat digunakan untuk mengukur kemampuan peserta didik yang lebih tinggi. Keunggulan tes pilihan ganda yaitu memiliki sifat yang lebih representatif dalam hal mencakup dan mewakili materi yang telah diajarkan pada peserta didik, serta memungkinkan guru untuk lebih objektif dalam mengoreksi lembar jawab, menentukan bobot skor, dan menentukan nilai hasil tes (Sudijono, 2006:135).

### **2.1.3.1 Kesahihan Instrumen Asesmen**

Kemampuan peserta didik mengikuti pembelajaran dapat diketahui melalui penggunaan instrumen asesmen sebagai alat evaluasi. Arikunto

(2013:40) mendeskripsikan alat sebagai sesuatu yang dapat digunakan untuk mempermudah seseorang dalam melaksanakan tugas atau mencapai tujuan secara lebih efektif dan efisien.

Arikunto (2013:40) menyatakan kata “instrumen” biasa disebut juga dengan kata “alat”. Dengan demikian, instrumen evaluasi juga dikenal dengan alat evaluasi. Karakteristik yang harus dimiliki oleh alat evaluasi yang baik menurut Sudijono (2006:40) adalah:

1. Valid, berasal dari kata bahasa arab “*Shahi*h” yang artinya tepat, benar, atau absah.
2. Reliabel, berasal dari kata bahasa arab “*Tsaabit*” yang artinya ajeg/kemantapan. Maksudnya adalah ketika alat evaluasi diujikan beberapa kali pada subyek yang sama hasilnya tetap sama.
3. Obyektif, berasal dari kata bahasa arab “*Maudhuu*’i” yang artinya disusun apa adanya.
4. Praktis, berasal dari kata bahasa arab “*’amalii*” yang artinya sederhana/lengkap.

Kesahihan instrumen asesmen terpenuhi apabila memenuhi 4 karakteristik utama yang telah dijelaskan. Instrumen mempunyai validitas eksternal apabila kriteria di dalam instrumen disusun berdasarkan fakta-fakta empiris yang telah ada di lapangan. Validitas eksternal diuji dengan cara membandingkan kriteria pada instrumen dengan fakta-fakta empiris yang terjadi di lapangan. Apabila terdapat kesamaan antara kriteria dalam instrumen dengan fakta di lapangan, maka dapat dinyatakan instrumen tersebut memiliki validitas eksternal yang tinggi (Sugiyono, 2015: 183).

Validitas empirik suatu produk instrumen asesmen yang dikembangkan, dievaluasi berdasarkan hasil uji coba produk. Setelah instrumen asesmen dinyatakan valid, maka diperlukan evaluasi mengenai karakteristik alat evaluasi yang obyektif dan praktis. Salah satu cara yang dapat digunakan untuk mengetahui karakteristik tersebut adalah melalui persebaran angket yang memberikan informasi respon guru dan peserta didik terhadap instrumen asesmen yang dikembangkan.

Karakteristik alat evaluasi yang baik kemudian dikembangkan dan diperjelas oleh Arifin (2012:65) menjadi 8 yaitu:

1. Valid, artinya suatu alat ukur dapat dikatakan valid jika betul-betul mengukur yang hendak diukur secara tepat.
2. Reliabel, artinya suatu alat ukur dapat dikatakan reliabel atau handal jika mempunyai hasil yang taat asas (*consistent*).
3. Relevan, artinya alat ukur yang digunakan harus sesuai dengan standar kompetensi, kompetensi dasar, dan indikator yang telah ditetapkan. Alat ukur juga harus sesuai dengan domain hasil belajar, seperti domain kognitif, afektif, dan psikomotor.
4. Representatif, artinya materi alat ukur harus betul-betul mewakili dari seluruh materi yang disampaikan. Hal ini dapat dilakukan bila guru menggunakan silabus sebagai acuan pemilihan materi tes. Guru juga harus memperhatikan proses seleksi materi, mana materi yang bersifat aplikatif dan mana yang tidak, mana yang penting dan mana yang tidak.
5. Praktis, artinya mudah digunakan. Kepraktisan ini bukan hanya dilihat dari sisi pembuat alat ukur, tetapi juga bagi orang lain yang ingin menggunakan alat ukur tersebut.
6. Deskriminatif, artinya adalah suatu alat ukur harus disusun sedemikian rupa, sehingga dapat menunjukkan perbedaan-perbedaan yang sekecil apapun. Semakin baik suatu alat ukur, maka semakin mampu alat ukur tersebut menunjukkan perbedaan secara teliti. Untuk mengetahui apakah suatu alat ukur cukup deskriminatif atau tidak, biasanya didasarkan atas uji daya pembeda alat ukur tersebut.
7. Spesifik, artinya suatu alat ukur disusun dan digunakan khusus untuk objek yang diukur. Jika alat ukur tersebut menggunakan tes, maka jawaban tes jangan menimbulkan ambivalensi atau spekulasi.
8. Proporsional, artinya suatu alat ukur harus memiliki tingkat kesulitan yang proporsional antara sulit, sedang, dan mudah.

Perolehan hasil penilaian yang akurat membutuhkan instrumen yang dapat mengungkapkan pengetahuan yang dimiliki peserta didik secara mendalam, mampu mengukur kemampuan berpikir peserta didik, dan dapat digunakan secara praktis. Arikunto (2013: 41-47) menggolongkan teknik evaluasi menjadi 2 macam yaitu, teknik nontes dan teknik tes. Teknik nontes tergolong menjadi beberapa macam diantaranya, skala bertingkat (*rating scale*), kuesioner (*questionair*), daftar cocok (*check list*), wawancara (*interview*), pengamatan (*observation*), dan riwayat hidup.

Teknik tes dibagi menjadi 3, yaitu tes diagnostik, tes formatif, dan tes sumatif. Tes diagnostik adalah tes yang digunakan untuk mengetahui kelemahan-kelemahan peserta didik sehingga berdasarkan hal tersebut dapat dilakukan penanganan yang tepat (Arikunto, 2013:48). Tes formatif adalah alat evaluasi formatif untuk mengetahui sejauh mana peserta didik telah terbentuk setelah mengikuti proses pembelajaran (Arikunto, 2013:50). Tes formatif diberikan pada akhir setiap proses pembelajaran. sedangkan tes sumatif adalah tes yang dilaksanakan setelah berakhirnya pemberian proses pembelajaran atau biasa dilaksanakan pada akhir semester (Arikunto, 2013:53).

### **2.1.3.2 HOTS**

*Higher Order Thinking Skills* (HOTS) disebut juga dengan kemampuan berpikir tingkat tinggi. HOTS adalah aktivitas berpikir peserta didik yang melibatkan level kognitif tingkat tinggi dari taksonomi berpikir Bloom meliputi menganalisis, mengevaluasi dan mencipta (Anderson & Krathwohl, 2001:43). Heong *et al.* (2011:121) juga mendefinisikan kemampuan berpikir tingkat tinggi adalah penggunaan pikiran secara luas yang melibatkan level kognitif hirarki tinggi dari taksonomi berpikir Bloom untuk menemukan tantangan baru. Kemampuan berpikir tingkat tinggi menuntut seseorang untuk menerapkan informasi baru atau pengetahuan yang telah dimilikinya dan memanipulasi informasi untuk menjangkau kemungkinan jawaban dalam situasi yang baru. Brookhart &



Llc (2011:5) menyebutkan tujuan pengajaran berdasarkan taksonomi kognitif Bloom menghendaki peserta didik untuk dapat menerapkan pengetahuan dan keterampilan untuk konteks baru, yaitu peserta didik dapat mengaplikasikan konsep yang belum terpikirkan sebelumnya.

Horizons *et al.*, (2016:158) mereview taksonomi Bloom yang telah direvisi, keterampilan berpikir pada ranah kognitif terbagi menjadi enam tingkatan, yaitu mengingat (*remember*), memahami (*understand*), aplikasi (*apply*), analisis (*analyze*), evaluasi (*evaluate*), dan menciptakan (*create*). Taksonomi Bloom yang telah direvisi kemampuan berpikir tingkat tinggi melibatkan analisis (C4), mengevaluasi (C5) dan mencipta (C6) dianggap berpikir tingkat tinggi. Tabel 2.4 menyajikan deskripsi kata kunci setiap kategori dalam Taksonomi Bloom terkait pembagian kemampuan berpikir.

Tabel 2.4 Dimensi Proses Berpikir

	Ranah Kognitif	Kata Kerja Operasional
HOTS	Mengkreasi (C6) <i>Mengkreasi ide/gagasan sendiri</i>	Mengkonstruksi, desain, kreasi, mengembangkan, menulis, memformulasikan
	Mengevaluasi (C5) <i>Mengambil keputusan sendiri</i>	Evaluasi, menilai, menyanggah, memutuskan, memilih, mendukung
	Menganalisis (C4) <i>Menspesifikasi aspek-aspek/elemen</i>	Membandingkan, memeriksa, mengkritisi, menguji
MOTS	Mengaplikasi (C3) <i>Menggunakan informasi pada domain berbeda</i>	Menggunakan, mendemonstrasikan, mengilustrasikan, mengoperasikan
	Memahami (C2) <i>Menjelaskan ide/konsep</i>	Menjelaskan, mengklasifikasi, menerima, melaporkan
LOTS	Mengetahui (C1) <i>Mengingat kembali</i>	Mengingat, mendaftar, mengulang, menirukan

(Narayanan *et al.*, 2015:3)

Widana (2017:3) mengkategorikan kemampuan berpikir tingkat tinggi menjadi 5 kemampuan yaitu, memecahkan masalah (*problem solving*), keterampilan berpikir kritis (*critical thinking*), berpikir kreatif (*creative thinking*), kemampuan berargumen (*reasoning*), dan kemampuan

mengambil keputusan (*decision making*). Item soal HOTS disusun berdasarkan pada aspek berpikir kognitif dimana sebagian besar memiliki pokok soal berupa pernyataan-pernyataan berisi kasus atau masalah yang terjadi di lingkungan peserta didik dan kehidupan sehari-hari untuk memberi stimulus dalam menyelesaikan soal tersebut. Penilaian yang mengukur HOTS dapat menggunakan bentuk tes subjektif dan tes objektif (Wardany *et al.*, 2017:2).

Pengembangan instrumen asesmen *testlet* ini disusun berdasarkan taksonomi Bloom secara berjenjang dimulai dari tingkatan berpikir yang paling rendah ke tinggi. Sehingga instrumen yang dikembangkan ini diharapkan mampu membantu peserta didik mengerjakan soal berpikir tingkat tinggi. Penyusunan instrumen asesmen berupa tes subjektif maupun tes objektif sebagai evaluasi dimudahkan dengan menggunakan kata operasional sebagaimana disajikan pada pada Tabel 2.4. Instrumen asesmen yang menjadi produk dalam penelitian ini yaitu alat evaluasi yang diambil dengan teknik tes berupa tes formatif, dalam bentuk tes objektif (pilihan ganda) berupa *testlet* berbasis STEM.

#### **2.1.4 Tinjauan Materi Tekanan Zat**

Tekanan merupakan suatu ukuran yang terdiri dari besarnya gaya yang bekerja pada suatu benda untuk setiap satu satuan luas permukaan bidang tekan. Tekanan dapat dinotasikan sebagai simbol P (*Pressure*). Satuan tekanan yang lain adalah pascal (Pa) atau  $\text{N/m}^2$ .

##### **2.1.4.1 Tekanan zat padat**

Tekanan pada suatu zat padat dapat dinyatakan sebagai gaya per satuan luas penampang. Secara matematis, tekanan dapat dinyatakan sebagai berikut.

$$P = \frac{F}{A}$$

Keterangan:

P = tekanan ( $\text{N/m}^2$ )

F = gaya (N)

$A = \text{luas bidang tekan (m}^2\text{)}$

#### 2.1.4.2 Tekanan zat cair

Tekanan pada zat cair sering disebut juga dengan tekanan hidrostatik. Tekanan hidrostatik ini tergantung pada suatu tingkatan kedalaman dan berat jenis pada zat cair. Tekanan pada zat cair mengarah ke segala arah. Rumus tekanan hidrostatik sebagai berikut:

$$P_h = \rho \times g \times h$$

Keterangan:

$P_h$  = tekanan hidrostatik zat cair (N/m<sup>2</sup>)

$\rho$  = massa jenis (kg/m<sup>3</sup>)

$g$  = percepatan gravitasi (m/s<sup>2</sup>)

$h$  = kedalaman dari permukaan (m)

Suatu benda yang dicelupkan ke dalam zat cair akan mendapat gaya angkat yang sebanding dengan volume zat cair yang dipindahkan benda itu. Sebuah benda yang dicelupkan sebagian atau seluruhnya akan mendapat gaya angkat oleh zat cair sebesar berat zat cair yang dipindahkan, hal ini merupakan bunyi dari hukum Archimedes. Alat – alat yang prinsip kerjanya berdasarkan hukum Archimedes antara lain pembuatan kapal laut, galangan kapal, kapal selam, balon udara. Secara matematis yang ada hubungan gaya apung dapat dituliskan sebagai berikut ini :

$$W' = W - F_a$$

Keterangan:

$W'$  = gaya berat benda di dalam air (N)

$W$  = gaya berat benda di udara (N)

$F_a$  = gaya apung atau gaya ke atas (N)

Gaya apung juga dapat dituliskan sebagai berikut ini :

$$F_a = \rho \times g \times V$$

Keterangan:

$F_a$  = gaya apung atau gaya ke atas (N)

$\rho$  = massa jenis (kg/m<sup>3</sup>)

$g$  = percepatan gravitasi ( $m/s^2$ )

$V$  = volume benda ( $m^3$ )

Tekanan dalam zat cair sangat berguna dalam kehidupan sehari-hari. Salah satu contohnya seperti yang dirumuskan oleh Pascal “Tekanan yang diberikan pada zat cair dalam ruang tertutup akan diteruskan kesegala arah dengan sama besar”. Banyak peralatan yang menggunakan prinsip Pascal antara lain dongkrak hidrolik, rem hidrolik, mesin pengangkat mobil hidrolik, dan kempa hidrolik. Secara matematis hukum pascal dapat dirumuskan sebagai berikut ini :

$$\frac{F_1}{A_1} = \frac{F_2}{A_2}$$

Keterangan:

$F_1$  = gaya pada tabung 1

$F_2$  = gaya pada tabung 1

$A_1$  = luas area pada tabung 1

$A_1$  = luas area pada tabung 1

#### 2.1.4.3 Tekanan gas

Semua zat memiliki massa dan menempati ruangan, tidak terkecuali zat gas. Hasil kali tekanan dengan volume suatu gas adalah tetap asal suhu zat tetap. Sebagai contoh adalah jika kita memompa ban sepeda, udara bisa masuk ke dalam ban jika pompa penghisap kita tekan, akhirnya udara masuk.

## 2.2 Kajian Penelitian yang Relevan

Berdasarkan uraian yang dijelaskan sebelumnya, ditemukan sejumlah hasil penelitian yang mendukung pengembangan instrumen asesmen. Rusilowati *et al.*, (2016:5719) telah mengembangkan instrumen asesmen yang digunakan sebagai strategi peningkatan kemampuan literasi sains yang rendah dimana menurut PISA (*Program for International Student Assessment*), Indonesia berada pada peringkat 60 dari 65 negara. Hasil evaluasi pengembangan instrumen tersebut memenuhi validitas isi, konstruk, dan penyelarasan dengan kriteria valid, dan sangat valid.

Penelitian lain tentang pengembangan instrumen asesmen juga dilakukan Okaviani *et al.*, (2015:326) yang memperoleh beberapa fakta bahwa: guru tidak menerapkan evaluasi pembelajaran dengan menyusun sendiri soal yang diujikan, tetapi mengadaptasi dari soal-soal yang ada di buku ajar/LKS; guru tidak membuat kisi-kisi sehingga ketercapaian hasil belajar yang diukur tidak jelas; dan peserta didik kesulitan ketika tipe soal yang diberikan guru saat ujian tidak sama dengan soal pada kegiatan belajar.

Kajian relevan juga dilakukan Shiell & Slepko (2015:201-210) dalam jurnal yang berjudul *integrated testlets: a new form of expert-student collaborative testing* yang merupakan artikel konseptual. Relevansi kajian yang ada dengan penelitian ini yaitu, jurnal Shiell & Slepko berkedudukan sebagai sumber pedoman penyusunan *testlet* secara struktural dan formatif. Dalam artikel ini mengungkapkan bahwa *testlet* diprediksi akan menjadi super-IT (*Integrated Testlet*) karena mampu mengintegrasikan model tes yang mengkonfirmasi langsung jawaban setiap langkah dalam menjawab soal secara signifikan. Inovasi model soal berupa *testlet* telah banyak dikembangkan dan diimplementasikan.

Penelitian lain tentang pengembangan *testlet* dilakukan oleh Nova *et al.*, (2016:1197-1203) dengan menyusun soal *testlet* model berjenjang yang masing-masing 4 soal pendukung menggunakan model ADDIE (*Analysis, Design, Develop, Implement, dan Evaluate*). Hasilnya menunjukkan bahwa 10 soal *testlet* yang dikembangkan nomor 1 dan 6 tidak valid sehingga soal tersebut tidak digunakan lagi untuk uji pemakaian. Yamtinah *et al.*, (2014:1-10) juga mengembangkan model *testlet* dengan 3 soal pendukung dengan desain penskoran menggunakan model GRM. Tujuan penelitiannya yaitu untuk mengetahui kemudahan dan kemanfaatan profil individu dalam mendeteksi kelemahan dan kelebihan peserta didik. Hasilnya menunjukkan bahwa *testlet* yang dikembangkan dilengkapi program analisis merupakan alternatif asesmen yang dapat digunakan guru.

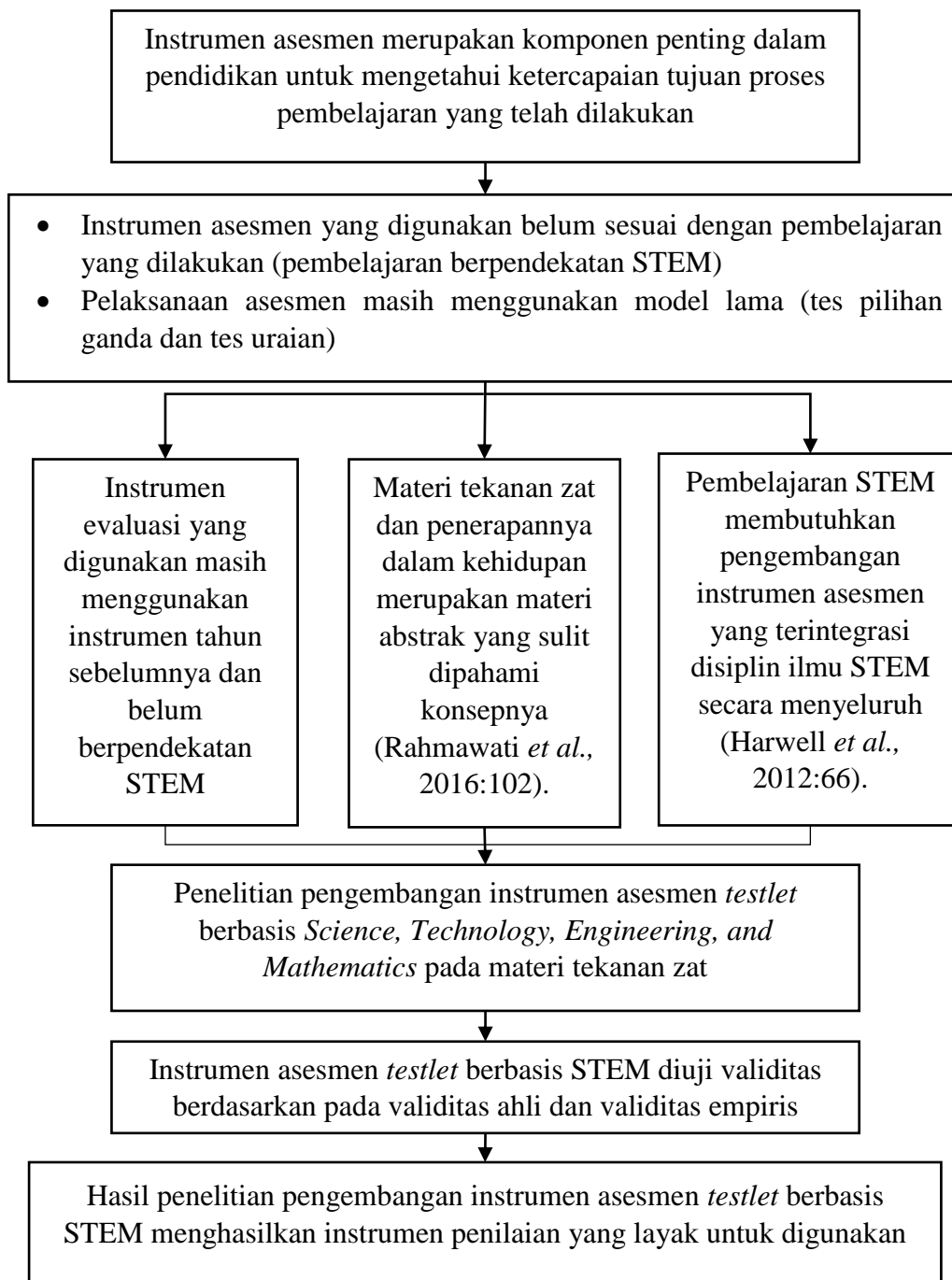
Penelitian Murti *et al.*, (2018:26-31) dalam jurnal berjudul “Studi Komparasi antara Tes Testlet dan Uraian dalam Mengukur Hasil Belajar Kognitif Siswa Kelas XI SMA Negeri 1 Gombang”. Penelitian ini bertujuan untuk mengetahui apakah tes testlet dapat berfungsi sebagaimana tes uraian. Penelitiannya merupakan penelitian eksploratif yang dilakukan untuk menggali suatu gejala yang relatif masih baru. Hasilnya diperoleh bahwa tes *testlet* dapat berfungsi sebagaimana tes uraian.

Pengembangan instrumen asesmen dengan berbagai macam model tes dilakukan untuk mengukur kemampuan peserta didik secara spesifik. Penelitian relevan telah dilakukan Winastiti *et al.*, (2018:37-42) dengan mengembangkan instrumen evaluasi berbasis literasi sains dalam jurnal yang berjudul *development of evaluation tool based on science literacy on topics of water recycling*. Tujuan penelitiannya yaitu untuk mengembangkan instrumen evaluasi berbasis literasi sains yang sesuai untuk diterapkan, dipahami dan mudah dipelajari oleh peserta didik. Hasil pengembangan instrumen evaluasi yang dikembangkan dinyatakan sangat valid pada bagian konstruk, isi, dan bahasa. Profil kemampuan literasi sains peserta didik yang diuji menggunakan instrumen evaluasi yang dikembangkan menunjukkan penguasaan literasi sains peserta didik kelas 5 SD Negeri 02 Gisikdrono masuk dalam kategori sedang. Relevansi dengan penelitian yang dilakukan Winastiti adalah sama-sama mengembangkan perangkat pembelajaran. Jika Winastiti mengembangkan instrumen evaluasi berbasis literasi sains, maka penelitian ini akan mengembangkan instrumen asesmen (evaluasi) berbasis STEM. Penyusunan asesmen berbasis STEM yang dilakukan peneliti didasarkan pada kebutuhan instrumen asesmen berbasis STEM dengan model *testlet*.

### **2.3 Kerangka Berpikir**

Kerangka berpikir disusun berdasarkan kesenjangan yang dengan harapan yang akan dicapai. Sejauh ini belum dikembangkan instrumen asesmen yang dapat digunakan sebagai instrumen asesmen

pembelajaran STEM. Pengembangan instrumen asesmen *testlet* STEM diharapkan dapat memenuhi kebutuhan instrumen asesmen kebutuhan di SMP Negeri 8 Salatiga. Bagan alur kerangka berpikir penelitian pengembangan ini selengkapnya ditunjukkan pada Gambar 2.3.



Gambar 2.3 Bagan Alur Kerangka Berpikir

Hasil dari penelitian ini diharapkan adanya instrumen asesmen *testlet* berbasis STEM yang valid dan reliabel dengan kompetensi penilaian pembelajaran IPA pada materi tekanan zat.



## BAB 5 SIMPULAN DAN SARAN

### 5.1 Simpulan

Berdasarkan pada hasil penelitian yang telah dilakukan, diperoleh simpulan sebagai berikut:

1. Instrumen asesmen *testlet* berbasis STEM pada materi tekanan zat memiliki validitas  $\geq 0,266$ , 10 *testlet* yang dikembangkan telah memenuhi kriteria valid, reliabilitas sebesar 0,866 dengan kategori reliabilitas sangat tinggi, analisis uji daya beda diperoleh 5 soal *testlet* dengan kategori cukup dan 5 soal *testlet* dengan kategori jelek, dan analisis indeks kesukaran diperoleh 2 soal dengan kategori mudah, 4 sedang, dan 4 sukar. *Testlet* yang telah dikembangkan dilengkapi dengan analisis untuk mengukur daya serap peserta didik sehingga dapat digunakan guru sebagai alternatif asesmen pendeteksi kesulitan belajar peserta didik. Dengan demikian *testlet* berbasis STEM memberi kemudahan guru dalam mengetahui kemampuan daya serap setiap peserta didik, efisiensi waktu mengkoreksi jawaban, memperoleh profil individu peserta didik, sekaligus mampu mendeteksi kesulitan belajar peserta didik.
2. Respon peserta didik dan guru berdasarkan persebaran angket terhadap pengembangan instrumen asesmen *testlet* berbasis STEM pada materi tekanan zat menunjukkan respon positif dengan rata-rata respon peserta didik 79,3% dan rata-rata respon guru 86%.

### 5.2 Saran

Saran yang dapat diberikan penulis untuk penelitian selanjutnya adalah sebagai berikut:

1. Soal yang akan digunakan dalam soal *testlet* sebaiknya dipersiapkan sebaik mungkin dan menyesuaikan dengan materi dan tingkat pemahaman peserta didik.
2. Diperlukan pengembangan instrumen yang dapat menjadi alternatif bagi guru dalam menjalankan peran ganda asesmen, selain sebagai pengukur

hasil belajar peserta didik juga sebagai alat memprediksi pemahaman konsep peserta didik.

3. Melengkapi instrumen *testlet* dengan program analisis data untuk mendeteksi kesulitan belajar peserta didik yang memunculkan profil peserta didik berupa indikator-indikator yang telah dikuasai dan belum dikuasai.

### DAFTAR PUSTAKA

- Afriana, J., Permanasari, A., & Fitriani, A. 2016. Penerapan Project Based Learning Terintegrasi STEM untuk Meningkatkan Literasi Sains Siswa Ditinjau dari Gender. *Jurnal Inovasi Pendidikan IPA*, 2(2), 202–212.
- Anderson, L.W., & Krathwohl, D.R. 2001. Kerangka Landasan untuk Pembelajaran, Pengajaran, dan Asesmen: Revisi Taksonomi Pendidikan Bloom (Terjemahan Agung Prihantoro). Yogyakarta: Pustaka Pelajar.
- Arifin, Z. 2012. *Evaluasi Pembelajaran (Edisi Revisi)*. Jakarta: Direktorat Jenderal Pendidikan Islam Kementerian Agama RI.
- Arikunto, S. 2013. *Dasar-Dasar Evaluasi Pendidikan*. Jakarta: Bumi Aksara.
- Asmalia, I., Fadiawati, N., & Nina, K. 2015. Pengembangan Instrumen Asesmen Berbasis Keterampilan Proses Sains pada Materi Stoikiometri. *Jurnal Pendidikan Dan Pembelajaran Kimia*, 4(1), 299–311.
- Asmawati, E. Y. S., Rosidin, U., & Abdurrahman. 2018. Efektivitas Instrumen Asesmen Model Creative Problem Solving pada Pembelajaran Fisika Terhadap Kemampuan Berpikir Kritis Siswa. *Jurnal Pendidikan Fisika*, 6(2), 128–143.
- Becker, K., & Park, K. 2011. Effects of Integrative Approaches 18 Among Science, Technology, Engineering, and Mathematics (STEM) Subjects on Students' Learning: A Preliminary Meta-analysis. *Journal of STEM Education*, 12(5), 23-37.
- Brookhart, S. M., & Llc, B. E. 2011. Educational Assessment Knowledge and Skills for Teachers. *Educational Measurement: Issues and Practice*, 30(1), 3–12.
- Brown, R., Brown J., Reardon K., & Merrill C. 2011. Understanding STEM: Current Perception. *The Technology and Engineering Teacher*, 70(6), 5-9.
- Bybee, R. W. 2013. *The Case for STE Education: Challenges and Opportunities*. Arlington: National Science Teachers Association (NSTA) Press.
- Council, N. R. 2012. *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas*. Washington, D.C.: The National Academies Press.
- Creswell, J. W. 2015. *Research Design Pendekatan Metode Kualitatif, Kuantitatif, dan Campuran Edisi 4*. Yogyakarta: Pustaka Pelajar.
- Effiong, U., Nkwol, N., & Nsungu, N.U. 2014. Test Type. students' achievement in senior secondary school physics and eradication of poverty ang hunger in Nigeria. *IOSR Journal of Research & Method in Education*, 4(1), 01-05.
- Gonzalez, H. B., & Kuenzi, J. J. 2012. *Science, Technology, Engineering, and Mathematics (STEM) Education: A Primer*. Congressional Research Service.

- Haksani. 2013. Pengembangan Perangkat Assessment Berbasis Keterampilan Generik Sains pada Mata Kuliah Praktikum Kimia Dasar Lanjut. *Jurnal Chemica*, 14(1): 27-37.
- Harwell, M., Moreno, M., Moore, T. J., Guzey, S. S., Phillips, A., & Roehrig, G. H. 2012. A Study of STEM Assessments in Engineering, Science, and Mathematics for Elementary and Middle School Students. *A Measurement Study of STEM Assessments*, 115(2), 66–74.
- Heong, Y. M., Othman, W. B., Yunus, J. Bin, Kiong, T. T., Hassan, R. Bin, Mohaffyza, M., & Mohamad, B. 2011. The Level of Marzano Higher Order Thinking Skills among Technical Education Students. *International Journal of Social Science and Humanity*, 1(2), 121–125.
- Horizons, S. E., Quest, T., & Dawn, T. 2016. A Taxonomy for Learning , Teaching, and Assessing : A Revision of Bloom's Taxonomy of Educational Objectives. *Phi Delta Kappa International*, 83(3), 154–159.
- Ismayani, A. 2016. Pengaruh Penerapan STEM Project-Based Learning terhadap Kreativitas Matematis Siswa SMK. *Indonesian Digital Journal of Mathematics and Education*, 3(4), 264–272.
- Kusumaningrum, L., Yamtinah, S., Nugroho, A., & Saputro, C. 2015. Pengembangan Instrumen Tes Diagnostik Kesulitan Belajar Kimia SMA Kelas XI Semester I Menggunakan Model Teslet. *Jurnal Pendidikan Kimia*, 4(4), 36–45.
- Malik, A., Rosidin, U., & Ertikanto, C. 2018. Pengembangan Instrumen Asesmen HOTS Fisika SMA Menggunakan Model Inkuiri Terbimbing. *Jurnal Lentera Pendidikan Pusat Penelitian LPPM UM METRO*, 3(1), 11–25.
- Mardapi, D. 2012. *Pengembangan dan Evaluasi Pendidikan*. Yogyakarta: Nuha Medika.
- Matondang, Z. 2009. Validitas dan Reliabilitas Suatu Instrumen Penelitian. *Jurnal Tabularasa PPS UNIMED*, 6(1): 87-97.
- Murti., Wiyanto., & Hartono. 2018. Studi Komparasi antara Tes Testlet dan Uraian dalam Mengukur Hasil Belajar Kognitif Siswa Kelas XI SMA Negeri 1 Gombong. *Unnes Physics Education Journal*, 7(1), 23–31.
- Narayanan, S., Nadu, T., & Adithan, M. 2015. Analysis of Question Papers in Engineering Courses with Respect to HOTS (Higher Order Thinking Skills). *American Journal of Engineering Education*, 6(1), 1–10.
- Nova, A. R., Parno., & Koes, S. 2016. Pengembangan Instrumen Asesmen Penguasaan Konsep Tes Testlet pada Materi Suhu dan Kalor. *Jurnal Pendidikan: Teori, Penelitian, Dan Pengembangan*, 1(6), 1197—1203.
- Nurhadi. 2009. *Pendekatan dalam Penilaian*. Jakarta: Pustaka Sinar Harapan.
- Okaviani, E., Fadiawati, N., & Kadaritna, N. 2015. Pengembangan Instrumen

- Asesmen Berbasis Keterampilan Proses Sains pada Materi Hukum-hukum Dasar Kimia. *Jurnal Pendidikan Dan Pembelajaran Kimia*, 4(1), 324–338.
- Rahmawati, I., Hidayat, A., & Rahayu, S. 2016. Penguasaan Konsep IPA Siswa SMP pada Materi Tekanan pada Zat Cair dan Aplikasinya. *Jurnal Pendidikan Sains*, 4(3), 102–112.
- Rahmawati, I. L., Hartono, & Nugroho, S. E. 2015. Pengembangan Asesmen Formatif untuk Meningkatkan Kemampuan Self Regulation Siswa pada Tema Suhu dan Perubahannya. *Unnes Science Education Journal*, 4(2), 842–850.
- Roberts, A. 2012. A Justification for STEM Education. *Technology and Engineering Teacher*, 2(1), 1–5.
- Rudyatmi, E. & Rusilowati, A. 2012. *Evaluasi Pembelajaran*. Semarang: FMIPA Unnes.
- Rusilowati, A., Kurniawati, L., & Nugroho, S. E. 2016. Developing an Instrument of Scientific Literacy Assessment on the Cycle Theme. *International Journal of Environmental & Science Education*, 11(12), 5718–5727.
- Scalise, K., & Wilson, M. 2007. *Bundle Models for Computerized Adaptive Testing in E-Learning Assessment* (In D. J. W). Oregon: GMAC Conference on Computerized Adaptive Testing.
- Shiell, R. C., & Slepko, A. D. 2015. Integrated Testlets: A New Form of Expert-Student Collaborative Testing. *Collected Essays on Learning and Teaching*, 8(2), 201–210.
- Stankous, N. V. 2016. Constructive Response Vs . Multiple-Choice Tests In Math : American Experience And Discussion ( Review ). *European Scientific Journal*, ISSN 1857, 308–316.
- Sudijono, A. 2006. *Pengantar Evaluasi Pendidikan*. Jakarta: Raja Gafindo Persada.
- Sudijono, A. 2014. *Pengantar Statistik Pendidikan* Jakarta: Raja Gafindo Persada.
- Sugiyono. 2017. *Metode Penelitian Kuantitatif, Kualitatif dan R&D*. Bandung: CV Alfabeta.
- Sujarwanto, & Rusilowati, A. 2015. Pengembangan Instrumen Performance Assessment Berpendekatan Scientific pada Tema Kalor dan Perpindahannya. *Unnes Science Education Journal*, 4(1), 780–787.
- Suprananta, K. 2012. *Pengukuran dan Penilaian Pendidik*. Yogyakarta: Graha Ilmu.
- Suryani, N. A., Sakti, I., & Purwanto, A. 2018. Perbandingan Hasil Belajar antara Model Pembelajaran CLIS (Children's Learning in Science) dengan Menggunakan Media KIT IPA di SMP Negeri 21 Kota Bengkulu. *Journal of*

- Science Education*, 2(1), 113–116.
- Susetyo, B. 2015. *Prosedur Penyusunan dan Analisis Tes untuk Penilaian Hasil Belajar Bidang Kognitif*. Bandung: PT Refika Aditama.
- Susongko, P. 2010. Perbandingan Keefektifan Bentuk Tes Uraian dan Testlet dengan Penerapan Graded Response Model (GRM). *Jurnal Penelitian Dan Evaluasi Pendidikan*, 14(2), 269–288.
- Trianto. 2013. *Model Pembelajaran Terpadu*. Jakarta: Bumi Aksara.
- Tsupros, N., Kohler, R., & Hallinen, J. 2009. *STEM education: A project to identify the missing components*. Pennsylvania: Intermediate Unit 1 and Carnegie Mellon.
- Wahidmurni., A. Mustikawan., & Ridho, A. 2010. *Evaluasi Pembelajaran: Kompetensi dan Praktik*. Yogyakarta: Nuha Letera.
- Wardany, K., Sajidan., & Ramli, M. 2017. Pengembangan Penilaian Untuk Mengukur Higher Order Thinking Skills Siswa. *Jurnal Inkuiri*, 6(2), 1–16.
- Widana, I. W. 2017. *Modul Penyusunan Soal Higher Order Thinking Skill (HOTS)*. Jakarta: Direktorat Jenderal Pendidikan Dasar dan Menengah.
- Winastiti, R. N., Rusilowati, A., & Prihatin, T. 2018. Development of Evaluation Tool Based on Science Literacy on Topics of Water Recycling. *Journal of Primary Education*, 8(1), 37–42.
- Yamtinah, S., Haryono., & Martini, K. S. 2014. Profil Individu Peserta Didik Pelengkap Tes Jenis Testlet sebagai Alternatif Pendeteksi Kesulitan Belajar Kimia. *Jurnal Profesi Pendidik*, 1(1), 1–10.
- Yusuf, M. 2015. *Asesmen dan Evaluasi Pendidikan*. Jakarta: Penamedia Group.