

Received : 22 May 2018  
Revised : 6 June 2018  
Accepted : 26 June 2018  
Published: 30 June 2018

DOI: doi.org/10.21009/1.04103

# The Development of High Order Thinking Skills (HOTS) Assessment Instrument for Temperature and Heat Learning

Aniq Rif'atun Najihah<sup>1, a)</sup>, Vina Serevina<sup>1, b)</sup>, Mutia Delina<sup>2, c)</sup>

<sup>1</sup>Program Studi Magister Pendidikan Fisika, FMIPA, Universitas Negeri Jakarta, Jakarta Timur 13220

<sup>2</sup>Program Studi Fisika, FMIPA, Universitas Negeri Jakarta, Jakarta Timur 13220

Email: <sup>a)</sup>aniqnajihah@yahoo.com, <sup>b)</sup>vina\_serevina@yahoo.com, <sup>c)</sup>mutia\_delina@unj.ac.id

## Abstract

This research aims to develop the High Order Thinking Skills (HOTS) test instrument. The instrument was presented into a valid and reliable questionnaire to get information of the grade-XI student's HOTS on temperature and heat. The research was done at Universitas Negeri Jakarta. The instrument has been trialed to 41 respondences on April 2018 at SMAN 1 Kabupaten Tangerang. The research was adapted from ADDIE approach which consists of Analyze, Design, Develop, Implement and Evaluate. The instrument consisted of 15 essays which the coefficient of reliability was 0.568, the difficulties and discrimination average were 0.817 and 0.363. Therefore the test instrument is proper to measure the student's HOTS.

**Keywords:** instrument, high order thinking skill, temperature, heat, ADDIE approach

## Abstrak

Penelitian ini bertujuan untuk mengembangkan instrumen penilaian *High Order Thinking Skills* (HOTS). Instrumen penilaian berupa soal tes yang valid dan reliabel untuk mengukur HOTS peserta didik pada materi suhu dan kalor kelas XI. Penelitian pengembangan ini dilakukan di Universitas Negeri Jakarta. Instrumen penilaian telah diujicoba pada bulan April 2018 di SMAN 1 Kabupaten Tangerang dengan 41 responden. Penelitian pengembangan ini diadaptasi dari model pengembangan ADDIE, yang meliputi: *Analyze* (Analisis), *Design* (Desain), *Develop* (Pengembangan), *Implement* (Penerapan), dan *Evaluate* (Evaluasi). Instrumen penilaian HOTS berupa soal tes yang terdiri dari 15 butir soal uraian dengan koefisien reliabilitas sebesar 0,568, rata-rata tingkat kesukaran 0,817 dan rata-rata daya pembeda 0,363. Jadi, Instrumen penilaian HOTS pada materi suhu dan kalor yang dikembangkan cukup mampu mengukur keterampilan *High Order Thinking Skills* peserta didik.

**Kata-kata kunci:** instrumen, *high order thinking skills*, suhu, kalor, pendekatan ADDIE

## PENDAHULUAN

Dalam proses pembelajaran, penilaian diperlukan untuk mengetahui tingkat pengetahuan peserta didik (Tungkasamit 2012; Arikunto 2012). Penilaian dapat berupa serangkaian pertanyaan yang harus dijawab atau perintah-perintah yang harus dikerjakan (Sudijono 2009). Agar penilaian yang

dilakukan efektif, maka perlu dikembangkan instrumen yang baku, valid, dan reliabel (Serevina 2016; Farra 2015). Selain itu faktor kemampuan pengajar dalam membuat instrumen penilaian yang sesuai dengan level kognitif peserta didik juga sangat diperlukan (Suzuk 2012).

Menurut taksonomi Bloom yang telah direvisi, terdapat enam level kognitif peserta didik; mengingat (C1), memahami (C2), menerapkan (C3), menganalisis (C4), mengevaluasi (C5) dan mengkreasi (C6). Mengingat, memahami, dan menerapkan termasuk dalam *Low Order Thinking Skill* (LOTS) atau berfikir tingkat rendah, sedangkan menganalisis, mengevaluasi, dan mengkreasi termasuk dalam *High Order Thinking Skill* (HOTS) atau berfikir tingkat tinggi (Yahya 2012; Saputra 2016). LOTS digunakan pada siswa sekolah dasar dan sekolah menengah pertama, sementara HOTS digunakan pada siswa sekolah menengah atas (SMA). Berikut ini adalah penjelasan HOTS pada peserta didik (King 2013; Kuseri 2014):

1. C4, menganalisis

Menganalisis merupakan mengolah informasi untuk memahami dan mencari hubungan dari informasi-informasi tersebut. Misalnya kegiatan membandingkan, mengorganisasi, menata ulang, mengajukan pertanyaan, dan menemukan.

2. C5, mengevaluasi

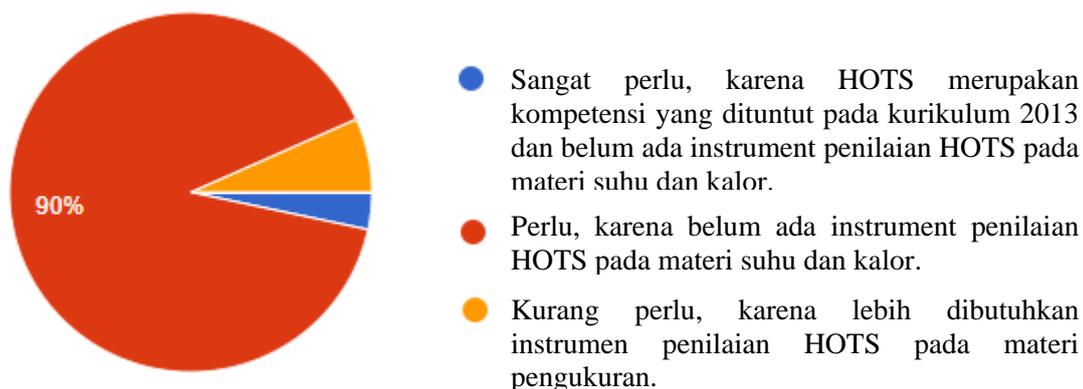
Mengevaluasi adalah menilai suatu keputusan atau tindakan. Contohnya pada kegiatan memeriksa, membuat hipotesa, mengkritik, bereksperimen, dan memberi penilaian.

3. C6, mengkreasi

Mengkreasi adalah menghasilkan ide-ide baru, produk, atau cara pandang terhadap sesuatu. Contohnya pada kegiatan mendesain, membangun, merencanakan, dan menemukan.

Dari penjelasan di atas, dapat disimpulkan bahwa HOTS menggabungkan berfikir kritis, reflektif, metakognitif, dan kreatif (Roets 2016).

Pada tahun 2015, Pratiwi melakukan pengembangan instrument penilaian HOTS berbasis kurikulum 2013 terhadap sikap disiplin peserta didik SMA (Pratiwi 2015). Penilaian HOTS tersebut sangat efektif untuk siswa yang sangat aktif, mandiri namun memiliki berkemampuan yang rendah dalam menyelesaikan soal fisika secara sistematis. Mengacu pada ide pengembangan yang telah dilakukan Pratiwi (2015), penelitian ini mengembangkan instrumen penilaian HOTS pada pelajaran fisika khususnya materi suhu dan kalor. Untuk mengetahui tingkat kebutuhan para guru maka telah dilakukan analisis kebutuhan instrumen penilaian HOTS pada materi suhu dan kalor kepada 30 orang guru SMA di Jakarta, Depok, Tangerang dan Bekasi. Dari analisis kebutuhan tersebut diperoleh data 3,3% guru atau 1 responden sangat memerlukan, 90% guru atau 27 orang responden memerlukan, dan 6,7% atau 2 responden tidak memerlukan instrumen penilaian HOTS pada materi suhu dan kalor. Hasil analisis kebutuhan tersebut ditunjukkan oleh GAMBAR 1 berikut ini.



**GAMBAR 1.** Hasil Analisis Kebutuhan yang dilakukan ke tiga puluh guru SMA di Jakarta, Depok, Tangerang, dan Bekasi

Berdasarkan analisis kebutuhan di atas, maka perlu dikembangkan instrumen penilaian HOTS berupa soal tes berbentuk uraian pada materi suhu dan kalor. Pengembangan instrumen penilaian HOTS ini bertujuan untuk menghasilkan instrumen penilaian HOTS yang valid dan reliabel untuk peserta didik.

## METODE PENELITIAN

Paragraf. Untuk menghasilkan instrumen penilaian HOTS, penelitian ini menggunakan *Research and Development Method* dengan pendekatan ADDIE yang terdiri dari lima langkah; *analyze, design, develop, implement* dan *evaluate* (Branch 2009). Penelitian dilakukan di Universitas Negeri Jakarta dan diujicoba kepada 41 responden di SMAN 1 Kabupaten Tangerang. Berikut ini merupakan langkah-langkah yang dilakukan dalam penelitian

### 1. Analisis Kebutuhan

Analisis kebutuhan merupakan tahapan awal penelitian untuk mendapatkan informasi potensi masalah yang ada. Pengumpulan informasi ini dilakukan dengan cara menyebar angket secara random kepada 30 orang guru SMA di Jakarta, Depok, Tangerang dan Bekasi.

### 2. Desain Instrumen

Tahap kedua adalah mendesain instrumen dengan menentukan kompetensi, bentuk instrumen dan materi bahan ajar.

#### a. Menentukan Kompetensi

Kompetensi yang digunakan dalam instrumen penilaian adalah Kompetensi Dasar (KD) 3.5 yaitu menganalisis pengaruh kalor dan perpindahan kalor yang meliputi karakteristik termal suatu bahan, kapasitas, dan konduktivitas kalor pada kehidupan sehari-hari. KD kemudian disesuaikan dengan level taksonomi HOTS.

#### b. Menentukan Bentuk Instrumen

Soal uraian HOTS dipilih sebagai bentuk instrumen penilaian awal.

#### c. Menyesuaian Materi Ajar (Suhu dan Kalor)

Materi ajar yang digunakan adalah suhu dan kalor.

### 3. Pengembangan Instrumen.

Selanjutnya melakukan pengembangan pendahuluan (*draft* awal) instrumen penilaian berdasarkan informasi analisis kebutuhan. Langkah awal yang dilakukan adalah menentukan atau membuat kisi-kisi (*blue print*) berupa matriks spesifikasi instrumen (indikator, pernyataan, nomor butir soal, dan skala penilaian). Selanjutnya menyusun kriteria acuan penilaian sesuai dengan butir penilaian. Kriteria-kriteria ini membantu guru dalam menilai peserta didik. Semakin banyak kriteria yang terpenuhi maka nilai HOTS peserta didik akan semakin tinggi.

### 4. Implementasi

Pada tahap implementasi dilakukan uji skala kecil dan uji skala besar

#### a. Uji Skala Kecil

Uji coba skala kecil dilakukan kepada tiga orang peserta didik untuk memperoleh informasi kualitas awal produk.

#### b. Revisi

Informasi kualitas awal produk dari uji skala kecil kemudian dianalisis. Hasil analisis tersebut digunakan untuk merevisi komponen instrumen penilaian yang belum valid.

#### c. Uji Skala Besar

Selanjutnya dilakukan uji skala besar kepada 41 orang peserta didik yang dipilih secara random.

#### d. Revisi

Data-data yang diperoleh dari uji skala besar digunakan untuk menghitung tingkat validitas dan reliabilitas instrumen penilaian. Kemudian produk direvisi sampai mendapatkan instrumen penilaian peserta didik berbasis HOTS yang valid dan reliabel, sehingga dapat digunakan secara masal.

### 5. Evaluasi

Evaluasi dapat dilakukan pada setiap tahap sebagai dasar untuk merevisi produk. Pada tahap ini kelayakan, validitas dan reabilitas instrumen penilaian diuji oleh beberapa ahli.

#### a. Uji Kelayakan Instrumen oleh Ahli

Uji kelayakan instrumen dilakukan dengan menggunakan lembar telaah untuk mengetahui kelayakan butir penilaian. Lembar telaah berisi pendapat dari para ahli dan guru yang kemudian digunakan untuk merevisi instrumen penilaian. Hasil telaah kemudian dianalisis secara deskriptif dengan menggunakan persamaan berikut ini

$$P_s = \frac{n}{N} \times 100\% \tag{1}$$

Dimana  $P_s$  merupakan persentase skor,  $n$  merupakan jumlah skor yang diperoleh dan  $N$  adalah jumlah skor maksimum

b. Uji Validitas dan Reabilitas Instrumen Secara Konstruk

Instrumen penilaian HOTS kemudian diuji coba dan diuji validitas. Uji validitas instrumen dilakukan dengan validasi konstruk meliputi aspek kognitif, aspek afektif, dan aspek psikomotorik. Kevalidan instrumen penilaian yang dikembangkan dapat dihitung dengan menggunakan teknik korelasi product moment berikut ini

$$r_{XY} = \frac{n \sum XY - (\sum X)(\sum Y)}{\sqrt{\{n \sum X^2 - (\sum X)^2\} \{n \sum Y^2 - (\sum Y)^2\}}} \tag{2}$$

Dimana  $r_{xy}$  merupakan koefisien korelasi antara variabel  $X$  dan  $Y$ ,  $N$  adalah jumlah responden dan  $X$  merupakan skor butir yang dicari validitasnya. Hasil uji validitas instrumen berbasis HOTS kemudian dibandingkan dengan  $r$ -tabel. Jika nilai signifikansinya 5% maka alat ukur dinyatakan valid. TABEL 1 menunjukkan interpretasi validitas soal menurut Arikunto (2012), TABEL 2 merupakan hasil uji validitas yang telah dilakukan pada produk.

**TABEL 1.** Interpretasi Validitas Soal (Arikunto 2012)

Rentang Nilai	Kategori
$0,8 < r_{XY} < 1,00$	Sangat Tinggi
$0,6 < r_{XY} < 0,8$	Tinggi
$0,4 < r_{XY} < 0,6$	Cukup
$0,2 < r_{XY} < 0,4$	Rendah
$0,00 < r_{XY} < 0,2$	Sangat Rendah

**TABEL 2.** Hasil Uji Validitas Produk

No Soal	$r_{hitung}$	$r_{tabel}$	Keterangan
1	-0,480	2,023	Drop
2	2,353	2,023	Valid
3	-0,523	2,023	Drop
4	7,585	2,023	Valid
5	4,396	2,023	Valid
6	2,809	2,023	Valid
7	2,437	2,023	Valid
8	1,104	2,023	Drop
9	2,974	2,023	Valid
10	8,037	2,023	Valid
11	3,545	2,023	Valid
12	3,306	2,023	Valid
13	2,032	2,023	Valid
14	0,910	2,023	Drop
15	0,024	2,023	Drop

Setelah instrumen berbasis HOTS dinyatakan valid maka kemudian instrumen diuji Reliabilitas. Reliabilitas menggambarkan konsistensi sebuah instrumen penilaian, artinya dalam dimensi apapun hasil yang diukur akan tetap sama. Tingkat reliabilitas instrumen penilaian dapat diukur dengan menggunakan persamaan Alpha Cronbach berikut

$$r_{11} = \left( \frac{n}{n-1} \right) \left( 1 - \frac{\sum \sigma_i^2}{\sigma_t^2} \right) \tag{3}$$

dimana,  $r_{11}$  adalah reliabilitas instrumen,  $\sum \sigma_i^2$  adalah jumlah varians skor tiap item,  $\sigma_t^2$  adalah varians total. Instrumen dikatakan bersifat reliabel apabila  $r_{11} > 0,6$  dengan taraf signifikansinya 5%. TABEL 3 berikut ini menunjukkan interpretasi reabilitas penilaian menurut Arikunto:

**TABEL 3.** Interpretasi Reliabilitas Penilaian (Arikunto 2012)

Rentang Nilai	Kategori
$0,8 < r_{xy} < 1,00$	Sangat Tinggi
$0,6 < r_{xy} < 0,8$	Tinggi
$0,4 < r_{xy} < 0,6$	Cukup
$0,2 < r_{xy} < 0,4$	Rendah
$0,00 < r_{xy} < 0,2$	Sangat Rendah

Kemudian daya pembeda soal diukur untuk membedakan antara peserta didik yang berkemampuan tinggi dengan peserta didik yang berkemampuan rendah. Perhitungan daya pembeda soal ditunjukkan oleh persamaan berikut ini.

$$D = \frac{B_A}{J_A} - \frac{B_B}{J_B} = P_A - P_B \tag{4}$$

dimana  $D$  adalah daya pembeda,  $JA$  adalah banyaknya peserta kelompok atas,  $JB$  adalah banyaknya peserta kelompok bawah,  $BA$  adalah banyaknya peserta kelompok atas yang menjawab soal dengan benar,  $BB$  adalah banyaknya peserta kelompok bawah yang menjawab soal dengan benar,  $PA$  adalah proporsi peserta kelompok atas yang menjawab benar, dan  $PB$  adalah proporsi peserta kelompok bawah yang menjawab benar. TABEL 4 berikut ini menunjukkan nilai interpretasi daya pembeda menurut Arikunto

**TABEL 4.** Interpretasi Daya Pembeda (Arikunto 2012)

Rentang Nilai (%)	Kategori
0-10	Sangat Buruk
11-19	Buruk
20-29	Sedang
30-49	Baik
50-100	Sangat Baik

Selain interpretasi daya pembeda, tingkat kesukaran soal juga diukur. Soal yang baik adalah soal yang tidak terlalu mudah atau tidak terlalu sukar. Soal yang terlalu mudah tidak merangsang peserta didik untuk meningkatkan usahanya saat memecahkan soal. Sebaliknya soal yang terlalu sukar akan menyebabkan peserta didik menjadi putus asa dan tidak mempunyai semangat untuk mencoba lagi. Bilangan yang menunjukkan mudah dan sukarnya suatu soal disebut indeks kesukaran dan dapat dihitung dengan persamaan berikut ini

$$P = \frac{B}{J_s} \tag{5}$$

dimana  $P$  adalah indeks kesukaran,  $B$  adalah banyaknya peserta didik yang menjawab soal dengan benar,  $J_s$  adalah jumlah seluruh peserta didik yang mengikuti penilaian.

**TABEL 5.** Interpretasi Tingkat Kesukaran Soal (Arikunto 2012)

Rentang Nilai (%)	Kategori
0-15	Sangat Sukar
16-30	Sukar
31-70	Sedang
71-85	Mudah
86-100	Sangat Mudah

TABEL 6 berikut ini menunjukkan hasil interpretasi tingkat kesukaran soal.

**TABEL 6.** Tingkat Kesukaran Soal

No. Soal	Tingkat Kesukaran	Kategori
1	0,077	Mudah
2	0,353	Sedang
3	0,083	Mudah
4	0,772	Mudah
5	0,576	Sedang
6	0,410	Sedang
7	0,364	Sedang
8	0,174	Sukar
9	0,430	Sedang
10	0,790	Sedang
11	0,494	Sedang
12	0,468	Sedang
13	0,309	Sedang
14	0,144	Sukar
15	0,004	Sangat sukar

Penilaian nilai hasil evaluasi instrumen selanjutnya dibandingkan dengan kriteria pada tabel berikut:

**TABEL 7.** Kriteria Evaluasi (Akbar 2013)

Persentase (%)	Kriteria
85-100	Sangat valid, dapat digunakan tanpa revisi
70-85	Cukup valid, dapat digunakan namun perlu direvisi
50-70	Kurang valid, disarankan tidak dipergunakan karena perlu revisi besar
0-50	Tidak valid, tidak boleh dipergunakan

## HASIL DAN PEMBAHASAN

Produk yang dikembangkan dalam penelitian ini adalah instrumen penilaian HOTS materi suhu dan kalor untuk peserta didik SMA kelas XI. Instrumen penilaian terdiri dari kisi-kisi instrumen HOTS materi suhu dan kalor dan 15 butir soal uraian. Validasi instrumen melibatkan dosen fisika Universitas Islam Syech Yusuf Tangerang, dosen fisika dasar Universitas Muhammadiyah Tangerang, guru fisika SMAN 9 Kabupaten Tangerang, guru fisika SMA Islamic Village dan guru fisika SMAN 1 Klari Karawang. Data yang diperoleh dari validasi pengembangan instrumen HOTS adalah sebagai berikut:

**TABEL 8.** Hasil Validasi Instrumen HOTS materi suhu dan kalor

No	Aspek yang dievaluasi	Validator					(%)	Kriteria Validasi	Ket.
		1	2	3	4	5			
1	Konstruksi butir instrument	76	78	80	85	78	79	Cukup valid, dapat digunakan namun perlu revisi	Revisi
2	Materi Suhu dan Kalor	85	81	80	85	81	82,8	Cukup valid, dapat digunakan namun perlu revisi	Revisi
3	Bahasa	84	85	80	88	85	82,8	Cukup valid, dapat digunakan namun perlu revisi	Revisi
Total							81,5		

TABEL 8 memperlihatkan bahwa persentase masing-masing kriteria instrumen penilaian HOTS materi suhu dan kalor sudah valid dan memerlukan revisi minor. Nilai presentase rata-rata validasi ahli adalah 81,5%, dengan demikian materi cukup valid untuk dipergunakan. Dengan demikian dapat disimpulkan bahwa secara keseluruhan pengembangan instrumen penilaian HOTS materi suhu dan

kalor cukup valid. Namun untuk penyempurnaan perlu dilakukan revisi pada konstruksi butir instrumen. Untuk menyempurnakan pengembangan instrumen penilaian HOTS perbaikan dilakukan berdasarkan saran dan komentar validator. Setelah divalidasi, hasil revisi pengembangan instrumen penilaian HOTS dinyatakan valid.

## SIMPULAN

TABEL 8 memperlihatkan bahwa persentase masing-masing kriteria instrumen penilaian HOTS materi suhu dan kalor sudah valid dan memerlukan revisi minor. Nilai presentase rata-rata validasi ahli adalah 81,5%, dengan demikian materi cukup valid untuk dipergunakan. Dengan demikian dapat disimpulkan bahwa secara keseluruhan pengembangan instrumen penilaian HOTS materi suhu dan kalor cukup valid. Namun untuk penyempurnaan perlu dilakukan revisi pada konstruksi butir instrumen. Untuk menyempurnakan pengembangan instrumen penilaian HOTS perbaikan dilakukan berdasarkan saran dan komentar validator. Setelah divalidasi, hasil revisi pengembangan instrumen penilaian HOTS dinyatakan valid.

## REFERENSI

- Akbar, S 2013, *Instrumen Perangkat Pembelajaran*, Remaja Rosdakarya Offset, Bandung.
- Arikunto, S 2012, *Dasar-dasar Evaluasi Pendidikan*, Bumi Aksara, Jakarta.
- Branch, RM 2009, *Instructional Design: The ADDIE Approach*, Springer Science & Business Media, LLC, New York.
- Farra, S 2015, 'Development of An Assessment Instrument to Evaluate Performance of The Skill of Decontamination', *Nurse Education Today*, pp. 1016.
- King, F, Goodson, L, & Rohani, F 2013, *Higher Order Thinking*, Florida State University, Florida.
- Kusaeri, 2014, *Acuan & Teknik Penilaian Proses & Hasil Belajar Kurikulum 2013*, Ar-Ruzz Media, Yogyakarta.
- Pratiwi, U, & Farida, E 2015, 'Pengembangan Instrumen Penilaian HOTS Berbasis Kurikulum 2013 Terhadap Sikap Disiplin', *Jurnal Penelitian dan Pembelajaran IPA*, pp. 123-142.
- Roets, L, Jeanette, & Maritz 2016, 'Facilitating the Development of Higher-Order Thinking Skills (HOTS) of Novice Nursing Postgraduates in Africa'. *Nurse Education Today*, p. 3.
- Saputra, H 2016, *Pengembangan Mutu Pendidikan Menuju Era Global*, CV. Smile's Indonesia Institute, Jakarta.
- Serevina, V, Dewi, C, & Mulyati, D 2016, 'Rancangan Tes dan Evaluasi Fisika yang Informatif dan Komunikatif pada Materi Kinematika Gerak Lurus'. *JPPPF-Jurnal Penelitian & Pengembangan Pendidikan Fisika*, vol. 2, no. 1, p. 81.
- Sudjiono, A 2009, *Pengantar Statistik Pendidikan*. Rajawali, Jakarta.
- Suzuk, E, & Ogan Bekiroglu, F 2012, 'Pre-service physics teachers intention toward classroom assessment', *ICEEPSY 2012*, p. 854.
- Tungkasmit, A, & Junpeng, P 2012, 'The Development of Authentic Assessment Training Curriculum for Research-Based Learning Class in Higher Education of Thailand', *International Conference on Education and Educational Psychology (ICEEPSY 2012)*, p. 1169.
- Yahya, A, Toukal, Z, & Osman, A 2012, 'Bloom's Taxonomy Based Classification for Item Bank Questions Using Support Vector Machines', *In Modern Advances in Intelligent Systems and Tools*, pp. 135-140.

