



**KEBEZAAN KEFUNGSIAN ITEM BERKAITAN JANTINA (GDIF) DALAM
INSTRUMEN TIPS II**

IZAZOL BINTI IDRIS



**TESIS YANG DIKEMUKAKAN UNTUK MEMENUHI SEBAHAGIAN
DARIPADA SYARAT MEMPEROLEH
SARJANA PENDIDIKAN**

**FAKULTI PENDIDIKAN
FAKULTI KEBANGSAAN MALAYSIA
BANGI**

2012



PENGAKUAN

Saya akui karya ini adalah hasil kerja saya sendiri kecuali nukilan dan ringkasan yang setiap satunya telah saya jelaskan sumbernya.

20 Februari 2012


IZAZOL BINTI IDRIS
P47666



PENGHARGAAN

Segala kesyukuran dipanjatkan kepada Allah swt kerana dengan izin dan kurniaannya, kajian dan laporan ini dapat diselesaikan. Kesempatan ini diguna untuk menzahirkan ucapan terima kasih kepada mereka yang terlibat secara langsung dan tidak langsung dalam kajian ini.

Setinggi-tinggi penghargaan dan ucapan terima kasih buat Kementerian Pengajian Tinggi Malaysia dan Universiti Pendidikan Sultan Idris kerana membiayai pengajian peringkat Sarjana ini. Terima kasih juga kepada Kementerian Pelajaran Malaysia, Jabatan Pelajaran Negeri, dan sekolah-sekolah yang terlibat, atas bantuan dan sokongan yang diberikan sepanjang kajian dilaksanakan.

Setinggi-tinggi penghargaan juga buat penyelia, Encik Zolkepli Haron, kerana kesabaran beliau membimbing dan memberi tunjuk ajar sepanjang kajian dan penulisan tesis ini dijalankan. Ucapan terima kasih juga ditujukan para pensyarah dan kakitangan Fakulti Pendidikan, Universiti Kebangsaan Malaysia, kerana sokongan dan bantuan yang diberikan sepanjang tempoh pengajian.

Akhir sekali, terima kasih kepada ibu bapa dan keluarga, serta teman-teman yang banyak memberikan sokongan emosi dan dorongan sepanjang tempoh pengajian yang dilalui.



ABSTRAK

Kajian ini bertujuan untuk mengesan kebezaan kefungisian item berkaitan jantina (GDIF) ke atas instrumen *Test for Integrated Science Process Skills II* (TIPS II). Prosedur analisis DIF berdasarkan model Rasch telah digunakan dalam kajian ini. Analisis DIF ialah satu kaedah statistik untuk mengenalpasti item-item yang menunjukkan ciri statistik (kefungisian) berbeza kepada dua tau lebih kumpulan pelajar yang sama kebolehan. Data diperoleh menggunakan instrumen TIPS II yang mengandungi 36 item objektif aneka pilihan. Item-item ini mengukur lima kemahiran proses sains bersepadu. Sampel kajian terdiri daripada 570 pelajar tingkatan empat dari lima buah sekolah menengah harian di sebuah daerah dalam negeri Perak. Analisis data secara deskriptif menggunakan perisian SPSS versi 17.0 manakala analisis DIF menggunakan perisian Winsteps versi 3.49. Kajian menunjukkan GDIF wujud dalam item-item TIPS II. 19 daripada 36 item didapati lebih mudah kepada pelajar lelaki dan 17 item lagi lebih mudah kepada pelajar perempuan. Terdapat empat item yang mempunyai saiz GDIF serius, melebihi 4.0 logit, serta signifikan pada tahap $p < 0.05$. Kajian juga menunjukkan kemungkinan terdapat sumbangan faktor-faktor demografi dalam fenomena kewujudan GDIF dalam TIPS II. Implikasi kajian menegaskan lagi kepentingan ciri keadilan item dalam meningkatkan kualiti sesebuah instrumen. Ciri keadilan ini juga perlu diberi perhatian dalam proses pembinaan instrumen pengukuran. Kajian ini juga menunjukkan item-item yang dibina berdasarkan kaedah CTT boleh dianalisis menggunakan kaedah moden berasaskan model Rasch, sekaligus membenarkan pendapat bahawa kaedah model Rasch ini tidak bertentangan, sebaliknya menyempurnakan kaedah CTT.

GENDER RELATED DIFFERENTIAL ITEM FUNCTIONING (GDIF) IN TIPS II

ABSTRACT

The purpose of the present study is to detect gender related differential item functioning (GDIF) in the Test for Integrated Science Process Skills II (TIPS II). The DIF analysis procedure used was based on Rasch model. This analysis is a statistical method for identifying items showing different statistical feature (functioning) towards two or more different groups of student with the equal ability. The instrument used is TIPS II which comprised 36 objective items, measuring five integrated science process skills. The sampel were 570 Form Four students from five schools in a district in Perak. Descriptive data analysis was done using SPSS version 17.0 and Winsteps version 3.49 was used for DIF analysis. The outcome of the study indicated 19 out of 36 items recognized easier to male and 17 items were easier to female. Four items were detected showing serious DIF, with DIF size over 4.0 logit, and were also significant at $p < 0.05$. This study also showed that there was several demography factors contributed to the detection of GDIF items in TIPS II. Therefore it is very important to ensure the fairness of an item in order to enhance an instrument's quality. This study also showed that an instrument that was developed based on CTT method can still be analyzed using Rasch-based method, proving that the modern Rasch-based method was not contradict the traditional CTT method, instead both two were in fact completed each other.

KANDUNGAN

	Halaman
PENGAKUAN	ii
PENGHARGAAN	iii
ABSTRAK	iv
ABSTRACT	v
KANDUNGAN	vi
SENARAI JADUAL	x
SENARAI ILUSTRASI	xii
SENARAI GRAF	xiii
SENARAI SINGKATAN	xiv

BAB I PENGENALAN

1.1	Pengenalan	1
1.2	Latar Belakang	1
1.3	Pernyataan Masalah	4
1.4	Tujuan Kajian	6
1.5	Objektif Kajian	6
1.6	Soalan Kajian	7
1.7	Kerangka Konsep Kajian	9
1.8	Kepentingan Kajian	10
1.9	Batasan Kajian	11
	1.9.1 Format item	11
	1.9.2 Sampel kajian	11
	1.9.3 Tempat kajian	11
	1.9.4 Kaedah pengesanan gdif	11
1.10	Definisi Istilah	12
	1.10.1 Kemahiran proses sains bersepadu	12
	1.10.2 Kebezaan kefungsian item	13
	1.10.3 Indeks DIF	14
1.11	Rumusan	14

BAB II TINJAUAN KEPUSTAKAAN

2.1	Pengenalan	15
2.2	Kemahiran Proses Sains Bersepadu	15
2.3	<i>Test of Integrated Science Process Skills II (TIPS II)</i>	17
2.4	Perbezaan Pencapaian Sains antara Jantina	18
2.5	Keadilan Ujian dan Item Bias	21
2.6	Kebezaan Kefungsian Item (DIF)	23
2.6.1	Takrifan	24
2.6.2	Sejarah dan perkembangan kajian dif	26
2.6.3	Prinsip dan andaian dalam kajian dif	29
2.6.4	Kriteria padanan dalam kajian dif	29
2.6.5	Kumpulan sampel dalam kajian dif	30
2.7	Pengesanan DIF dalam Analisis Item	31
2.7.1	Teori Respons Item (IRT)	31
2.7.2	DIF dalam IRT	35
2.7.3	Model pengukuran Rasch	35
2.8	Kajian Lepas Mengenai GDIF	37

2.9	Rumusan	39
-----	---------	----

BAB III METODOLOGI

3.1	Pengenalan	40
3.2	Rekabentuk Kajian	40
3.3	Sampel Kajian	41
3.4	Instrumen Kajian	42
3.5	Tatacara Pemerolehan Data	45
3.6	Tatacara Analisis Data	46
3.6.1	Analisis peringkat I: Analisis awal prestasi sampel dan instrumen kajian	46
3.6.2	Analisis peringkat II: Analisis GDIF purata	50
3.6.3	Analisis peringkat III: Analisis GDIF mengikut tahap kebolehan	51
3.6.4	Analisis peringkat IV: Analisis GDIF untuk perbandingan keupayaan menjawab item-item TIPS II berdasarkan faktor bangsa dan aliran persekolahan	52
3.7	Rumusan	52

BAB IV DAPATAN KAJIAN

4.1	Pengenalan	53
4.2	Profil Sampel	53
4.3	Analisis Peringkat I: Analisis Awal Prestasi Sampel dan Prestasi Instrumen Kajian	54
4.3.1	Analisis prestasi sampel dalam instrumen TIPS II	54
4.3.1.1	Analisis prestasi keseluruhan sampel	55
4.3.1.2	Analisis prestasi sampel mengikut jantina	56
4.3.2	Analisis prestasi item dalam instrumen TIPS II	58
4.4	Analisis GDIF Purata	64
4.5	Analisis GDIF mengikut Tahap Kebolehan	69
4.5.1	Analisis GDIF pada tahap kebolehan tinggi	69
4.5.2	Analisis GDIF pada tahap kebolehan rendah	71
4.5.3	Corak GDIF dalam 36 Item TIPS II	72
4.6	Ringkasan Analisis GDIF dalam 36 Item TIPS II	73
4.7	Penilaian Distraktor bagi Item GDIF	75
4.7.1	Item 9	77
4.7.2	Item 16	79
4.7.3	Item 22	80
4.7.4	Item 29	81
4.8	Kesan Item GDIF ke atas Pengukuran Kebolehan Sampel	82
4.9	Perkaitan GDIF dengan Faktor Demografi	88
4.9.1	Analisis perbandingan keupayaan berdasarkan bangsa	89
4.9.2	Analisis perbandingan keupayaan berdasarkan aliran persekolahan	98
4.10	Rumusan	107

BAB V PERBINCANGAN

5.1	Pengenalan	108
5.2	Ringkasan Dapatan Kajian	108
5.3	Perbincangan	111
5.3.1	Prestasi pelajar dan kualiti item TIPS II	111
5.3.2	Kewujudan dan intensiti GDIF dalam TIPS II	112
5.3.3	Analisis GDIF pada dua tahap kebolehan	115
5.3.4	Kesan Item GDIF ke atas pengukuran kemahiran pelajar	116
5.3.5	Analisis GDIF berdasarkan bidang kemahiran	118

5.3.6	Analisis GDIF berdasarkan faktor demografi bangsa	120
5.3.7	Analisis GDIF berdasarkan faktor demografi aliran persekolahan	122
5.4	Implikasi dan Cadangan	123
5.5	Penutup	126
RUJUKAN		127
LAMPIRAN		
A	Surat Kebenaran Menjalankan Kajian daripada EPRD	132
B	Surat Kebenaran Menjalankan Kajian daripada JPN Perak	133
C	Analisis GDIF pada Tahap Kebolehan Tinggi	134
D	Analisis GDIF pada Tahap Kebolehan Rendah	136
E	Corak GDIF pada 36 Item TIPS II	138

**SENARAI JADUAL**

No. Jadual		Halaman
3.1	Taburan item ujian mengikut bidang kemahiran	43
3.2	Item dan objektif kemahiran yang diukur	43
3.3	Ciri-ciri TIPS II	45
3.4	Pengkategorian tahap kebolehan	51
4.1	Profil sampel	54
4.2a	Ringkasan prestasi keseluruhan sampel (1)	55
4.2b	Ringkasan prestasi keseluruhan sampel (2)	55
4.3a	Ringkasan prestasi sampel mengikut jantina (1)	57
4.3b	Ringkasan prestasi sampel mengikut jantina (2)	57
4.4	Dapatan ujian-t	58
4.5	Analisis item ujian TIPS II	59
4.6	Analisis GDIF ke atas 36 Item TIPS II	66
4.7	Rumusan analisis GDIF 36 Item TIPS II	68
4.8	Rumusan analisis GDIF pada kumpulan pelajar berkebolehan tinggi	70
4.9	Rumusan analisis GDIF pada kumpulan pelajar berkebolehan rendah	72
4.10	Kewujudan GDIF dalam 36 Item TIPS II	74
4.11	Ciri item GDIF	75
4.12	Analisis pilihan jawapan bagi item 9	78
4.13	Analisis pilihan jawapan bagi item 16	79
4.14	Analisis pilihan jawapan bagi item 22	80
4.15	Analisis pilihan jawapan bagi item 29	82
4.16	Ciri item GDIF untuk analisis pengecualian	83
4.17a	Analisis Ukuran kebolehan responden berdasarkan 36 item	83
4.17b	Analisis ukuran kebolehan responden berdasarkan 32 item	84
4.18a	Analisis ukuran kebolehan responden berdasarkan 36 item mengikut jantina	85





4.18b	Analisis ukuran kebolehan responden berdasarkan 32 item mengikut jantina	86
4.19	Analisis GDIF berdasarkan bangsa	91
4.20	Analisis GDIF berdasarkan aliran persekolahan	100



SENARAI RAJAH

No. Rajah		Halaman
1.1	Kerangka konsep kajian	9
2.1	Keluk Ciri Item	33
2.2	DIF seragam	34
2.3	DIF tak seragam	34
4.1	Peta Item-Individu	63
4.2	Item 9	77
4.3	Item 16	79
4.4	Item 22	80
4.5	Item 29	81

**SENARAI GRAF**

No. Graf		Halaman
4.1a	Perbandingan keupayaan menjawab item TIPS II berdasarkan kemahiran antara kategori bangsa (Lelaki)	93
4.1b	Perbandingan keupayaan menjawab item TIPS II berdasarkan kemahiran antara kategori bangsa (Perempuan)	94
4.1c	Perbandingan keupayaan menjawab item TIPS II berdasarkan kemahiran antara kategori jantina (Melayu)	95
4.1d	Perbandingan keupayaan menjawab item TIPS II berdasarkan kemahiran antara kategori jantina (Bukan Melayu)	96
4.2a	Perbandingan keupayaan menjawab item TIPS II berdasarkan kemahiran antara kategori aliran persekolahan (Lelaki)	102
4.2b	Perbandingan keupayaan menjawab item TIPS II berdasarkan kemahiran antara kategori aliran persekolahan (Perempuan)	103
4.2c	Perbandingan keupayaan menjawab item TIPS II berdasarkan kemahiran antara kategori jantina (Kebangsaan)	104
4.2d	Perbandingan keupayaan menjawab item TIPS II berdasarkan kemahiran antara kategori jantina (Jenis Kebangsaan)	105



SENARAI SINGKATAN

CTT	Classical Test Theory
BPPDP	Bahagian Perancangan dan Penyelidikan Dasar Pendidikan
DIF	Differential Item Functioning
ETS	Educational Testing Services
GDIF	Gender Differential Item Functioning
ICC	Item Characteristic Curve
IRT	Item Response Theory
JPN	Jabatan Pelajaran Negeri
KPM	Kementerian Pelajaran Malaysia
LPM	Lembaga Peperiksaan Malaysia
PBS	Pentaksiran Berasaskan Sekolah
PEKA	Pentaksiran Kerja Amali
PPK	Pusat Perkembangan Kurikulum
TIPS II	Test of Integrated Science Process Skills II
UKM	Universiti Kebangsaan Malaysia



BAB I

PENDAHULUAN

1.1 PENGENALAN

Bab ini menjelaskan kajian yang dijalankan untuk meninjau Kebezaan Kefungsian Item Berkaitan Jantina (GDIF) dalam item-item instrumen pengukuran kemahiran proses sains bersepadu, iaitu *Test of Integrated Process Skills* (TIPS II). Perbincangan menyentuh mengenai latar belakang kajian dan pernyataan masalah yang mencetuskan kajian ini. Kajian ini juga secara ringkasnya digambarkan dengan kerangka konsep dan diperjelaskan oleh kepentingan kajian, batasan kajian, dan definisi istilah yang digunakan.

1.2 LATAR BELAKANG

Pendidikan sains di Malaysia dipacu ke arah pemupukan budaya sains dan teknologi dengan memfokuskan pembangunan individu yang dinamik, teguh, mampu bersaing dan bertahan, dalam menguasai pengetahuan sains dan teknologi. Sehubungan itu, kurikulum sains kebangsaan digubal untuk mencapai matlamat ini dengan memberikan peluang kepada para pelajar untuk memperoleh pengetahuan dan kemahiran saintifik. Kurikulum sains yang digubal bertujuan melahirkan pelajar yang aktif, dan mereka diberi peluang untuk memerhati, bertanya soalan, membentuk dan menguji hipotesis, menganalisa, menakrif data, melapor, dan menilai dapatan. Pelajar digalakkan untuk mengaplikasi kemahiran proses sains dalam aktiviti pembelajaran mereka.





Kemahiran proses sains merupakan satu proses mental yang menggalakkan pemikiran kritis, kreatif, analitik, dan sistematik. Penguasaan kemahiran proses sains, bersama dengan sikap dan pengetahuan yang sesuai menjamin keupayaan pelajar untuk berfikir secara berkesan. Kemahiran proses sains terdiri daripada kemahiran proses asas dan kemahiran proses bersepadu. Kemahiran proses sains bersepadu disenaraikan oleh Pusat Perkembangan Kurikulum (2001) seperti berikut:

1. Menggunakan perhubungan ruang dan masa
2. Mentafsir maklumat
3. Mendefinisi secara operasi
4. Mengawal pembolehubah
5. Membuat hipotesis
6. Mengekspерimen

Penilaian melalui prosedur *hands-on* bagi menentukan kemahiran proses sains yang telah dikuasai pelajar dianggap paling sesuai (Ong, Zurida & Fong 2006). Dalam kurikulum pendidikan sains di Malaysia, prosedur penilaian ini telah dilaksanakan secara kerja amali, dikenali sebagai Pentaksiran Kerja Amali (PEKA), diperkenalkan pada tahun 1999. PEKA merupakan salah satu pentaksiran berasaskan sekolah (PBS), yang digubal oleh Kementerian Pelajaran Malaysia (KPM), untuk menilai tahap penguasaan pelajar dalam kemahiran proses sains. PEKA dipacu berdasarkan aktiviti *hands-on* dan *student directed research*. Dengan itu, ia diakui menilai secara lebih dekat sampel sebenar bagi setiap komponen kemahiran.

Diakui PEKA membolehkan guru mengetahui tahap penguasaan kemahiran proses sains pelajar (LPM, 2001). Maklum balas ini seterusnya membolehkan pelajar memperbaiki penguasaan kemahiran proses sains. Walau bagaimanapun, beberapa kajian dapat mengesan masalah dan kekangan dalam pelaksanaan PEKA. Antaranya kegagalan pelajar untuk mengenalpasti tujuan, menghubungkan konsep dan teori pengajaran yang diperolehi dengan eksperimen PEKA yang dijalankan (Nordin 1997). Selain itu penaksiran kemahiran sains juga memakan masa dan memerlukan pemerhati atau penaksir terlatih (Ong, Zurida & Fong 2006; Kazeni 2005). Masalah-masalah





berkaitan penaksiran ini boleh membawa kepada kemungkinan masalah kesukaran menentukan skor dengan tepat dan pemberian skor yang tidak adil.

Sehubungan itu, para pengkaji lepas seperti Ong, Zurida dan Fong (2006) dan Kazeni (2005) telah mencadangkan penggunaan ujian kertas dan pensel sebagai satu kaedah pengukuran kemahiran proses sains yang bersifat objektif dan boleh ditadbir secara efisien, sebagai alternatif kepada penaksiran berbentuk aktiviti *hands-on* di makmal. Disebabkan keperluan terhadap pengukuran yang piawai dan berkualiti rendah juga menyebabkan banyak instrumen pengukuran dibina melalui format item aneka pilihan (Hamilton & Snow 1988).

Beberapa instrumen sedemikian yang telah dibina dan digunakan secara meluas. Antaranya ialah *Test for Integrated Process Skill II* (TIPS II) oleh Burns, Wise dan Okey (1985). Mereka turut menyenaraikan beberapa instrumen lain yang dibina untuk tujuan sama seperti *Test of Integrated Process Skills* (TIPS) dan *Test of Integrated Science Processes* (TISP).



Dalam proses pembinaan instrumen, aspek penting yang sering dititik beratkan ialah kesahan dan kebolehpercayaan instrumen tersebut untuk mengukur konstruk yang hendak diukur. Selain itu, aspek lain yang kini menjadi fokus pembina instrumen ialah keadilan item. Walaupun aspek ini telah lama difokuskan oleh institusi penaksiran di luar negara seperti *Educational Testing Service* (ETS), Amerika Syarikat, namun isu ini masih baru dalam konteks pengujian dan penaksiran di Malaysia (Aminawati Zahara 2008).

Keadilan dalam sesuatu item instrumen adalah penting bagi memastikan tiada ralat pengukuran untuk menentukan keberkesanan instrumen tersebut. Item yang adil akan menunjukkan kefungsiannya yang sama kepada dua atau tiga kumpulan yang mempunyai perbezaan ciri atau latar belakang seperti jantina, bangsa, lokasi dan jenis sekolah, ataupun latar sosioekonomi. Item-item yang adil dapat menjamin kualiti instrumen yang terbina olehnya, untuk digunakan dalam sesuatu aktiviti pengujian. Dapatan daripada pengujian oleh item-item yang adil seterusnya menjamin penilaian selanjutnya dapat dilakukan dengan telus.





1.3 PERNYATAAN MASALAH

Beberapa kajian mengenai pencapaian dan kebolehan sains menunjukkan perbezaan antara lelaki dan perempuan. Klein et al. (1997) mendapati pelajar perempuan lebih cenderung untuk memperoleh skor yang rendah berbanding pelajar lelaki dalam ujian sains secara bertulis.

O'Neil dan Brown (1998) menegaskan dapatan dan kesimpulan yang dibuat berkaitan perbezaan keupayaan antara jantina ini boleh juga dikaji berdasarkan faktor berkaitan pelajar (bangsa, bahasa ibunda, dan lain-lain) dan faktor berkaitan ujian (format item, jenis item, dan sebagainya). Cadangannya ini menunjukkan faktor item dan ujian juga boleh mempengaruhi ukuran kebolehan seseorang pelajar.

Apabila satu instrumen mengukur pencapaian sains, dan dapatan menunjukkan pelajar lelaki dan perempuan, yang mempunyai kebolehan setara, adalah berbeza secara signifikan dalam memberi respons yang betul terhadap sesuatu item dalam instrumen tersebut, terdapat kemungkinan item itu bersifat bias atau berat sebelah terhadap satu kumpulan calon (sama ada kumpulan lelaki atau perempuan). Zumbo (2007) juga menekankan antara implikasi yang mungkin berlaku dengan kehadiran item bias dalam sesuatu instrumen ialah kesannya ke atas hubungan, beza min antara kumpulan, dan kesahan konstruk instrumen yang digunakan.

Secara empirikal, istilah kebezaan kefungsiian item (*Differential Item Functioning* - DIF) akan digunakan bagi merujuk kepada item bias. DIF wujud dalam sesuatu item apabila pelajar-pelajar yang mempunyai kebolehan yang sama, tetapi daripada kumpulan yang berasingan atau berbeza latar belakang, mempunyai kebarangkalian berbeza untuk mendapat jawapan yang betul. Item-item DIF ditentukan apabila item-item tersebut berfungsi secara berbeza bagi kumpulan calon yang berbeza latar belakang tetapi setara dari segi kebolehan yang hendak diukur. Sekiranya keputusan menunjukkan item-item tersebut memberi peluang yang tidak sama rata atau berat sebelah kepada calon dari kumpulan yang berbeza, maka kita katakan item itu berbeza dari segi kefungsiannya untuk mengukur konstruk terbahit (Siti Rahayah 2008). Apabila kebezaan kefungsiian ini difokuskan kepada jantina,





istilah *Gender Differential Item Functioning* (GDIF) digunakan. Istilah ini diterjemahkan ke Bahasa Melayu sebagai Kebezaan Kefungsian Item berkaitan Jantina.

Melalui analisis DIF, Hamilton dan Snow (1988) mendapati pelajar lelaki lebih berkebolehan menjawab item-item sains yang melibatkan pemprosesan visual dan pengetahuan yang melibatkan pengalaman mereka dalam kehidupan seharian di luar persekolahan. Kenyataannya ini selari dengan kenyataan Erickson dan Erickson (1984) bahawa pelajar menunjukkan prestasi yang lebih baik dalam item-item yang menggunakan objek atau peristiwa dalam kelompok “sfera pengalaman” mereka sendiri.

Sementara itu, sehubungan dengan kajian DIF yang dijalankannya ke atas item-item sains, Marshall (1991) telah mencadangkan kajian GDIF turut dilaksanakan ke atas item-item yang mengukur kemahiran sains. Ini bertepatan dengan keadaan di Malaysia, yang menekankan kebolehan seseorang pelajar dalam sains bukan hanya diukur berdasarkan pengetahuan sains sahaja, tetapi juga meliputi kemahirannya dalam proses sains. Penggubalan kurikulum sains yang menekankan penguasaan kemahiran proses sains telah membawa kepada keperluan terhadap satu instrumen yang sah dan boleh dipercayai untuk mengukur kemahiran ini.

Instrumen TIPS II yang telah dibina oleh Burns, Wise dan Okey (1985) telah digunakan di Malaysia oleh beberapa pengkaji seperti Ong, Zurida, dan Fong (1996), Lay (1999) dan Yeam (2007).

Terjemahan TIPS II ke dalam Bahasa Melayu oleh Ong, Zurida, dan Fong untuk digunakan dalam kajian mereka memberikan dapatan yang menunjukkan tahap penguasaan kemahiran proses sains dalam kalangan pelajar sekolah menengah adalah tidak memuaskan. Faktor yang menyumbang kepada dapatan sebegitu ialah pada masa kajian tersebut dijalankan, pengajaran kemahiran proses sains dalam bilik darjah masih belum dilaksanakan secara eksplisit (Ong, Zurida & Fong 2006).



Manakala TIPS II terus digunakan oleh pengkaji lain seperti Lay (1999) dan Yeam (2007) dalam kajian mereka yang berkaitan. Kedua-duanya menggunakan TIPS II untuk mengukur kemahiran proses sains dalam kalangan guru sekolah dan guru pelatih di IPG. Item-item TIPS II yang bersifat bebas kandungan kurikulum membolehkan ia digunakan di pelbagai peringkat usia. Ciri ini juga membolehkan pentadbiran TIPS II lebih meluas, bukan terhad kepada pelajar aliran sains sahaja. Hal ini menarik pengkaji untuk menganalisis item-item TIPS II bagi mengenalpasti kesesuaiannya untuk digunakan dalam kalangan pelajar sekolah menengah di Malaysia. Analisis ini memberi fokus kepada kebezaan kefungsiian item-item TIPS II berkaitan jantina.

1.4 TUJUAN KAJIAN

Kajian ini merupakan kajian rintis ke atas satu instrumen pengukuran kemahiran proses sains bersepadu, iaitu *Test for Integrated Science Process Skills II* (TIPS II) yang dibangunkan oleh Burns, Wise dan Okey (1985). Instrumen ini dipilih untuk kajian kerana ia adalah antara instrumen yang sering digunakan untuk mengukur kemahiran proses sains bersepadu dalam kalangan pelajar sekolah menengah mahupun guru pelatih dan guru di Malaysia. Fokus kajian adalah terhadap kewujudan item-item TIPS II yang menunjukkan kefungsiian yang berbeza berkaitan jantina (GDIF).

1.5 OBJEKTIF KAJIAN

Analisis DIF memberikan tumpuan kepada setiap item dalam TIPS II. Setiap item dikaji berdasarkan jawapan pelajar untuk melihat sama ada item tersebut menunjukkan kefungsiian berbeza terhadap kumpulan jantina pelajar. Analisis GDIF digunakan untuk mencapai objektif kajian berikut:

1. Mengesan kewujudan GDIF dalam item-item TIPS II.
2. Mengetahui intensiti GDIF dalam item-item TIPS II.
3. Mengenalpasti sama ada faktor bangsa boleh dikaitkan dengan kebezaan keupayaan pelajar lelaki dan perempuan menjawab item TIPS II.



4. Mengenalpasti sama ada aliran persekolahan boleh dikaitkan dengan kebezaan keupayaan pelajar lelaki dan perempuan menjawab item TIPS II.

1.6 SOALAN KAJIAN

Untuk memenuhi objektif kajian, beberapa persoalan kajian perlu dijawab. Soalan kajian yang perlu dijawab bagi memenuhi objektif pertama dan ke dua ialah:

1. Adakah terdapat GDIF dalam item TIPS II?
2. Apakah magnitud (saiz dan arah) GDIF dalam item TIPS II?
3. Berapakah bilangan dan peratusan item yang menunjukkan tahap GDIF signifikan dan penting dalam TIPS II?
4. Bagaimanakah corak GDIF dalam item-item yang dikaji?
5. Adakah item-item yang menunjukkan GDIF signifikan memberi kesan kepada kebolehan pelajar?
6. Adakah item tertentu lebih mudah kepada pelajar lelaki atau pelajar perempuan daripada latar belakang bangsa berbeza?
7. Adakah item tertentu lebih mudah kepada pelajar lelaki atau pelajar perempuan daripada latar belakang aliran persekolahan berbeza?

1.7 KERANGKA KONSEP KAJIAN

TIPS II telah dibangunkan oleh Burns, Wise, dan Okey pada tahun 1985 untuk mengukur kemahiran proses sains bersepadu pelajar. Instrumen ini mengukur lima kemahiran proses sains bersepadu, iaitu mengenalpasti pembolehubah, mengenalpasti dan menyatakan hipotesis, mendefinisi secara operasi, merekabentuk eksperimen, dan melukis graf dan mentafsir maklumat.

Prosedur analisis DIF dalam kajian ini berasaskan Model Rasch, dengan menggunakan perisian Winsteps versi 3.49. Prosedur ini digunakan untuk menganalisis setiap item dalam instrumen TIPS II bagi melihat sama ada wujud kebezaan kesukaran apabila ditadbir kepada kumpulan calon yang berbeza jantina tetapi setara dari segi kebolehan. Ciri statistik item yang dikaji untuk melihat GDIF





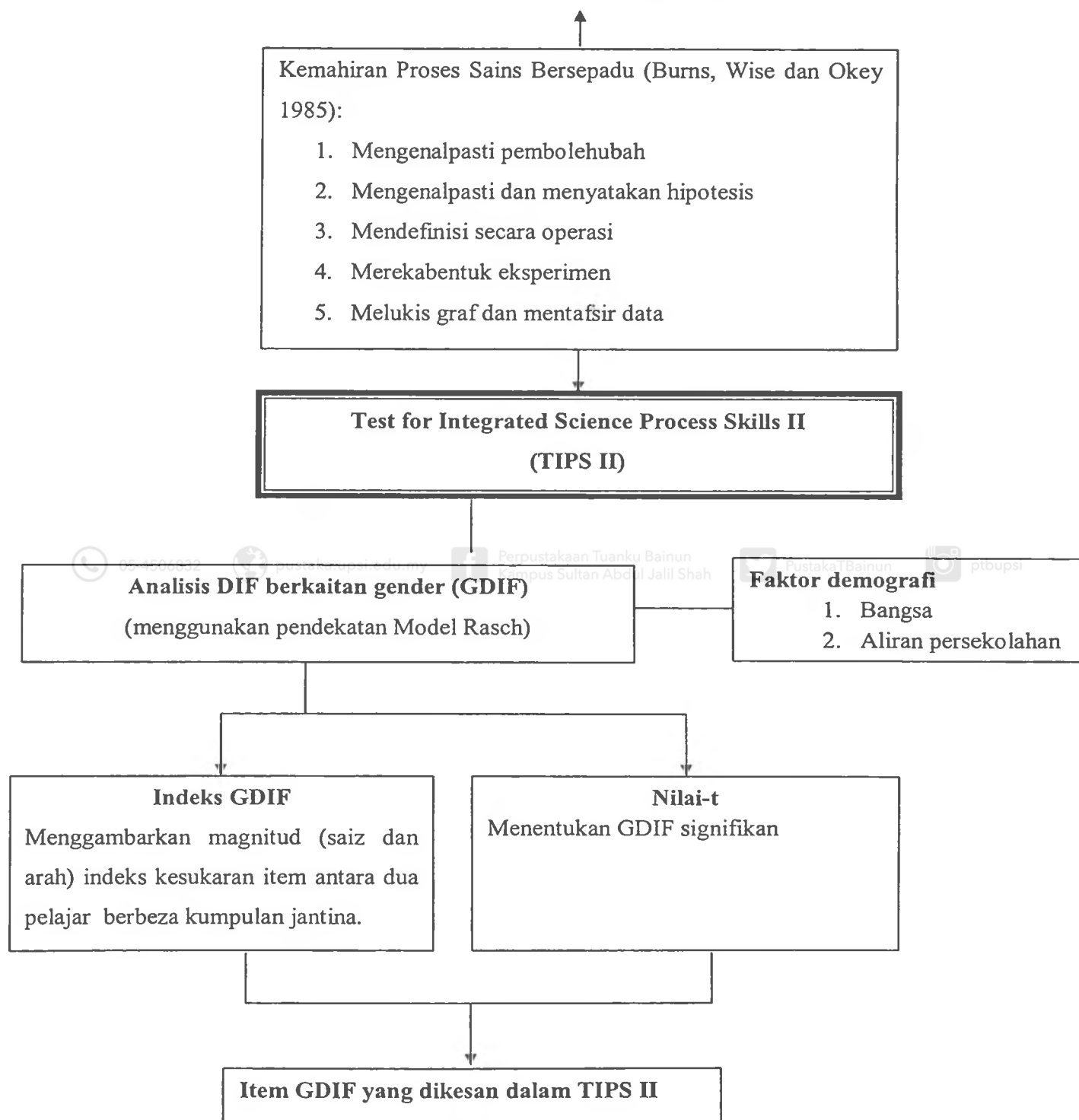
ialah indeks kesukaran item, yang ditentukan daripada jawapan pelajar. Seterusnya, indeks kesukaran item ini dibandingkan antara dua kumpulan jantina yang telah dipadankan mengikut tahap kebolehan mereka. Analisis secara perbandingan ini menghasilkan indeks GDIF, yang menunjukkan saiz perbezaan kesukaran item bagi dua kumpulan jantina tersebut (Hanizah 2007).

Indeks GDIF menunjukkan saiz dan arah kesukaran item yang dianalisis. Saiz GDIF menunjukkan perbezaan kesukaran item antara calon pelajar yang dibandingkan. Simbol positif dan negatif yang menandakan arah DIF pula menunjukkan kumpulan pelajar mana yang mendapat kelebihan apabila menjawab item tersebut. Seterusnya, nilai t-kiraan yang dipaparkan juga perlu dibandingkan dengan nilai t-kritikal = 2.0 bagi menentukan tahap kesignifikanannya.

Faktor demografi turut digambarkan dalam kerangka konsep bagi mengukuhkan dapatan analisis GDIF ke atas item-item TIPS II. Fokus diberikan kepada faktor bangsa dan aliran persekolahan pelajar. Dua faktor ini dikenalpasti berdasarkan kajian oleh Klein et al. (1997) iaitu faktor bangsa (etnik) dan ciri sekolah boleh mempengaruhi pencapaian pelajar dalam sains. Sehubungan dengan itu, kebezaan keupayaan pelajar lelaki dan perempuan dari latar belakang bangsa dan aliran berbeza ini untuk menjawab item-item TIPS II turut dikenalpasti.



Kebezaan keupayaan pelajar lelaki dan perempuan yang sama kebolehan menjawab item-item kemahiran proses sains dengan betul



Rajah 1.1 Kerangka konsep kajian

1.8 KEPENTINGAN KAJIAN

Kajian ini diharap dapat memberikan pendedahan kepada penyelidik dan pihak yang terlibat dalam pembinaan ujian agar mengambil berat mengenai aspek keadilan dalam item-item yang dibina. Kepelbagaian pelajar di Malaysia melibatkan latar belakang bangsa (kaum/etnik), bahasa ibunda, pendidikan ibu bapa, jenis dan aliran persekolahan, serta status sosioekonomi keluarga. Apabila populasi pelajar yang menduduki sesuatu ujian boleh dikategorikan kepada beberapa kumpulan kecil mengikut ciri tertentu, adalah penting bagi pembina ujian untuk melihat sama ada item yang digunakan itu adil kepada kumpulan pelajar yang sama kebolehan untuk menjawab dengan betul, tanpa mengira perbezaan jantina, bangsa, bahasa ibunda, dan sebagainya. Sekiranya ada kumpulan yang menanggapi item tertentu lebih sukar, maka item itu dikatakan berbeza dari segi kefungsiannya (menunjukkan DIF) dan perlu dikaji semula (Hanizah 2008).

Instrumen TIPS II adalah antara alat pengukuran kemahiran proses sains yang telah dibina dan diuji kesahannya di negara Barat. Namun kesahan, kebolehpercayaan, dan aspek keadilan item bagi instrumen ini belum dikenalpasti di Malaysia. Ramai pendidik menjalankan penyelidikan dengan mengandaikan kesahan dan kebolehpercayaan instrumen yang dibangunkan di negara Barat adalah sama untuk konteks Malaysia. Kajian ini dapat menentukan salah satu aspek penting dalam kesahan instrumen iaitu kefungsiannya item ke atas responden yang berbeza jantina.

Selain daripada itu, TIPS II yang telah dibangunkan pada 1985 menggunakan kaedah tradisional yang berdasarkan pendekatan Teori Ujian Klasik (*Classical Test Theory* – CTT). Analisis GDIF pula menggunakan model Rasch yang berdasarkan pendekatan moden Teori Respon Item (*Item Response Theory* – IRT). Sehubungan itu, selain daripada menjelaskan prosedur analisis item bagi mengesan item GDIF, kajian ini juga diharap dapat menunjukkan keserasian dua model pengukuran ini.

1.9 BATASAN KAJIAN

Batasan masa, kewangan, dan sumber mempengaruhi kebolehlaksanaan kajian. Ini seterusnya mempengaruhi keluasaan skop kajian. Batasan kajian ini disenaraikan seperti berikut:

1.9.1 Format Item

TIPS II mengandungi 36 item, dengan lima bidang kemahiran yang diukur. Walau bagaimanapun setiap kemahiran diwakili oleh bilangan item yang tidak sama. Item-item dipersembahkan dalam bentuk teks dan rajah, namun persembahan rajah (graf) hanya terhad kepada item-item bagi satu kemahiran sahaja. Kesemua item berbentuk aneka pilihan dengan empat opsyen. Instrumen ini juga ditadbir dalam Bahasa Inggeris.

1.9.2 Sampel Kajian

Kebezaan keupayaan menjawab item TIPS II hanya dilihat berdasarkan jantina. Manakala faktor demografi yang dikaitkan dengan item GDIF ialah bangsa dan aliran persekolahan. Faktor pembeza lain seperti bahasa ibunda, status sosioekonomi keluarga, dan lokasi sekolah tidak dikaji.

1.9.3 Tempat Kajian

Kajian ini hanya dijalankan di sebuah daerah dalam negeri Perak, melibatkan lima buah sekolah sahaja.

1.9.4 Kaedah Pengesanan GDIF

Kaedah pengesanan GDIF dalam kajian ini menggunakan perisian Winstep versi 3.49. Analisis yang diperolehi diringkaskan secara automatik oleh perisian dalam bentuk jadual yang menunjukkan aras kesukaran item, perbezaan aras kesukaran item, dan

nilai ujian-t bagi kumpulan pelajar lelaki dan perempuan yang dibandingkan. Kesignifikanan DIF dilihat berdasarkan ujian-t pada tahap $p < 0.05$.

1.10 DEFINISI ISTILAH

Beberapa istilah penting yang digunakan dalam kajian ini diperjelaskan seperti berikut:

1.10.1 Kemahiran Proses Sains Bersepadu

Kemahiran proses sains merupakan komponen kemahiran saintifik. Ia diperlukan untuk mencari jawapan kepada sesuatu masalah atau membuat keputusan secara bersistem. Kemahiran proses sains dikelaskan kepada dua, iaitu kemahiran asas dan kemahiran bersepadu.

Kemahiran proses sains bersepadu ialah kesinambungan kepada kemahiran proses sains asas yang sepatutnya diterapkan seawal pelajar berada di peringkat sekolah rendah. PPK (2001) menyenaraikan lima kemahiran proses sains bersepadu yang mesti dikuasai pelajar, iaitu kemahiran mentafsir data, mendefinisi secara operasi, mengawal pembolehubah, membina hipotesis, dan mengeksperimen. Kelima-lima kemahiran ini turut digunakan dalam TIPS II sebagai lima konstruk yang diukur. Penguasaan kemahiran proses sains berdasarkan TIPS II adalah apabila pelajar berupaya:

1. Mengenalpasti pembolehubah

Mengenalpasti pembolehubah bersandar, tidak bersandar, dan dimalarkan, berdasarkan kepada gambaran ringkas mengenai sesuatu eksperimen yang diberi.

2. Mengenalpasti dan menyatakan hipotesis

Memilih satu hipotesis yang boleh diuji berdasarkan gambaran mengenai tiga jenis pembolehubah yang terlibat dalam sesuatu eksperimen.

3. Mendefinisi secara operasi

Memberikan definisi operasi yang sesuai untuk setiap pembolehubah yang terlibat dalam sesuatu eksperimen yang digambarkan.

4. Merekabentuk eksperimen

Memilih satu rekabentuk eksperimen yang sesuai untuk menguji satu hipotesis atau masalah yang diberi.

5. Melukis graf dan mentafsir data

Mengenalpasti satu graf yang dapat mewakili data dan memerihalkan hubungan antara pembolehubah terlibat, berdasarkan kepada gambaran ringkas mengenai eksperimen yang dibuat dan data yang diperolehi.

1.10.2 Kebezaan Kefungsian Item

Kebezaan Kefungsian Item dikenali sebagai *Differential Item Functioning* (DIF) dalam Bahasa Inggeris, berlaku apabila individu dari kumpulan berbeza (kebiasaannya jantina atau bangsa) dengan trait pendam (kebolehan atau kemahiran) yang sama, mempunyai kebarangkalian berbeza untuk memberikan respons dalam sesuatu soal selidik atau memberikan jawapan betul dalam sesuatu ujian (Siti Rahayah 2008).

Analisis DIF menyediakan petunjuk tingkah laku yang tidak diduga oleh sesuatu item dalam ujian. Satu yang perlu dipastikan dengan jelas ialah, sesuatu item tidak menunjukkan DIF sekiranya individu dari kumpulan berbeza mempunyai kebarangkalian berbeza untuk memberikan respons, tetapi item tersebut akan menunjukkan DIF sekiranya individu dari kumpulan berbeza, dengan kebolehan sebenar yang setara, mempunyai kebarangkalian berbeza untuk memberikan respons atau jawapan. Secara lebih tepat, satu item menunjukkan DIF apabila didapati aras kesukaran (b), diskriminasi (a) atau asimptot lebih rendah (c) bagi item tersebut berbeza merentasi kumpulan calon yang menjawabnya. Nilai a, b, dan c dianggar berdasarkan Model Respons Item (*Item Response Theory* – IRT).



1.10.3 Indeks DIF

Menurut Hamilton dan Snow (1998), indeks DIF menunjukkan sama ada ahli dua kumpulan berbeza, dipadankan mengikut tahap kebolehan yang relevan, mempunyai kebarangkalian berbeza untuk menjawab item dengan betul. Indeks DIF digunakan dalam kajian ini untuk melihat saiz perbezaan kesukaran dan arah kecenderungan bagi setiap item yang dikaji ke atas kumpulan pelajar berdasarkan jantina.

1.11 RUMUSAN

Bab ini membincangkan perkara-perkara yang menjadi asas kepada kajian yang dilakukan. Kandungan bab ini disusun bermula dengan latar belakang yang menjelaskan betapa pentingnya item-item dalam sesuatu instrumen dibina dengan berkualiti, bukan sahaja dari aspek kesahan dan kebolehpercayaan, tetapi juga aspek keadilan. Seterusnya perbincangan dibawa kepada masalah kajian, objektif, dan persoalan kajian. Manakala kerangka konsep kajian menggambarkan tumpuan kajian ke atas analisis item dan melihat perbezaan pelajar dari kumpulan berlainan jantina tetapi mempunyai kebolehan yang setara, dalam menjawab item-item TIPS II. Kepentingan kajian juga dibincangkan sebagai justifikasi perlunya kajian ini dijalankan. Akhir sekali batasan kajian dan definisi istilah juga dibincangkan bagi memperjelaskan gambaran awal tentang kajian yang dijalankan.

