

MENGESAN KEBERBEZAAN FUNGSI ITEM BAGI ITEM-ITEM UJIAN PENCAPAIAN PRINSIP PERAKAUNAN TINGKATAN EMPAT DI NEGERI PERAK

*Lim Hooi Lian & Subashini A/P Sreenivasan

*hllim@usm.my

ABSTRAK

Kajian ini bertujuan mengesan keberbezaan fungsi item (Differential Item Functioning, DIF) bagi item-item ujian pencapaian Prinsip Perakaunan Tingkatan Empat di negeri Perak. Instrumen kajian merupakan satu set ujian pencapaian mata pelajaran Prinsip Perakaunan yang mengandungi 40 soalan aneka pilihan dengan empat fungsi pengganggu. Sampel kajian terdiri daripada 1000 orang pelajar dari 20 buah sekolah negeri Perak. Analisis data telah dijalankan dengan menggunakan Program Winsteps, Microsoft Excel, dan SPSS. Kaedah yang digunakan untuk mengesan DIF ialah analisa statistik menggunakan nilai $t \geq \pm 1.96$ pada $p = .05$. Saiz DIF ditentukan melalui nilai DIF Contrast $\pm .5$. Data dianalisis dan dapatan kajian menunjukkan: (1) Item-item di dalam ujian pencapaian Prinsip Perakaunan Tingkatan 4 menunjukkan 17 item sebagai item suspek DIF. Berdasarkan Nilai DIF Contrast, tiada item yang disahkan mempunyai saiz DIF yang besar dan boleh diabaikan; (2) Item-item di dalam ujian pencapaian Prinsip Perakaunan Tingkatan 4 menunjukkan lapan item sebagai item suspek DIF terhadap jantina. Berdasarkan nilai DIF Contrast, tiada item yang disahkan mempunyai saiz DIF yang besar dan boleh diabaikan. Gambaran yang komprehensif berasaskan maklumat perbezaan prestasi pelajar dengan item-item yang dikaji menunjukkan item tiada kesan DIF yang nyata.

Kata Kunci: *Differential Item Functioning (DIF); Prinsip Perakaunan; Program Winsteps*

PENGENALAN

Mata pelajaran Prinsip Perakaunan mendukung matlamat Falsafah Pendidikan Kebangsaan dalam melahirkan individu yang mempunyai akauntabiliti, berfikiran kritikal dan reflektif, mengamalkan budaya kerja yang beretika, berkemahiran dalam teknologi maklumat, membudayakan pembelajaran sepanjang hayat, mempunyai kemahiran komunikasi yang baik melalui pendekatan pembelajaran bermakna dan kesepaduan teori serta amalan perakaunan (Bahagian Pembangunan Kurikulum, 2015). Kurikulum pendidikan perakaunan kita memikul arah tuju dan objektif yang jelas dan unggul, selaras dengan matlamat Falsafah Pendidikan Kebangsaan (FPK) dalam menghadapi era globalisasi abad ke-21 yang semakin mencabar (Bahagian Pembangunan Kurikulum, 2015). Segala pengetahuan dan kemahiran semasa ini boleh diperolehi di peringkat sekolah agar memastikan individu mendapatkan peluang pekerjaan yang lebih luas dan baik (Hanushek & Kimko, 2000). Sehubungan itu, mata pelajaran Prinsip Perakaunan adalah induk kepada sistem pendidikan kita, iaitu amalan pengajaran dan pembelajaran mata pelajaran ini patutlah lebih mementingkan aspek penyediaan pelajar yang peka dan kritikal kepada perubahan global, terutamanya dalam bidang perdagangan dan ekonomi, agar berupaya menjadi tenaga kerja yang setanding dengan ekspektasi taraf antarabangsa dalam semua bidang kerjaya. Ekoran daripada itu, dapat melahirkan tenaga buruh yang berkemahiran tinggi atau professional (Bahagian Pembangunan Kurikulum, 2015).

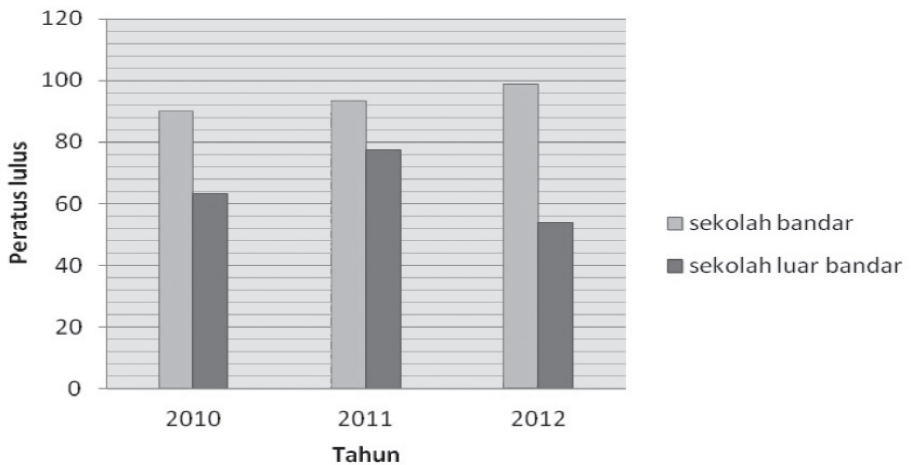
Dalam bidang pendidikan, ujian pencapaian mengukur tingkah laku pengajaran dan pembelajaran kognitif, afektif, dan psikomotor (Siti Rahayah Ariffin, 2003). Secara khususnya, ujian pencapaian diadakan untuk menilai prestasi dan kemampuan pelajar dalam sesuatu bidang. Ujian pencapaian untuk penguasaan pengetahuan sesuatu mata pelajaran seperti Prinsip Perakaunan diadakan untuk mengukur kebolehan pelajar dalam mata pelajaran tersebut (Zuraini et al., 2009). Selain itu, ujian pencapaian piawai diadakan secara selaras di seluruh negara iaitu peperiksaan Sijil Pelajaran Malaysia (SPM) meliputi sukatan kandungan Tingkatan Empat dan Lima. Ujian ini yang menentukan takdir pelajar di Malaysia untuk menempatkan pelajar meneruskan penerokaan pendidikan lanjutan ke institusi pengajian tinggi demi melayakkan diri sebagai tenaga kerja yang berkemahiran tinggi atau profesional. Maksy dan Lin Zheng (2008) telah mendapati bahawa gred yang dijangka akan didapati oleh pelajar dalam ujian pencapaian mempunyai hubungan yang signifikan ke atas performa mereka.

Berikutan itu, pengukuran yang berkualiti dalam mengukur pencapaian pelajar diperlukan khususnya dalam peperiksaan awam. Oleh itu, untuk memastikan bahawa item-item yang terdapat dalam sesuatu ujian tidak memihak kepada pihak yang tertentu perlu dititikberatkan. Kemungkinan boleh berlaku keadaan di mana sesetengah item dalam ujian menunjukkan '*Differential Item Functioning*' (DIF) iaitu item berfungsi berbeza-beza terhadap sesuatu

kumpulan (Ercikan, 2002). Perbezaan ini menunjukkan pelajar-pelajar yang mempunyai kebolehan yang sama, tetapi daripada kelompok kumpulan yang berlainan mempunyai kebarangkalian yang berbeza untuk mendapat jawapan yang betul (Linarce, 2013). Dengan ini, pembinaan kertas soalan untuk ujian tersebut perlu berkualiti dan adil untuk semua pelajar di Malaysia.

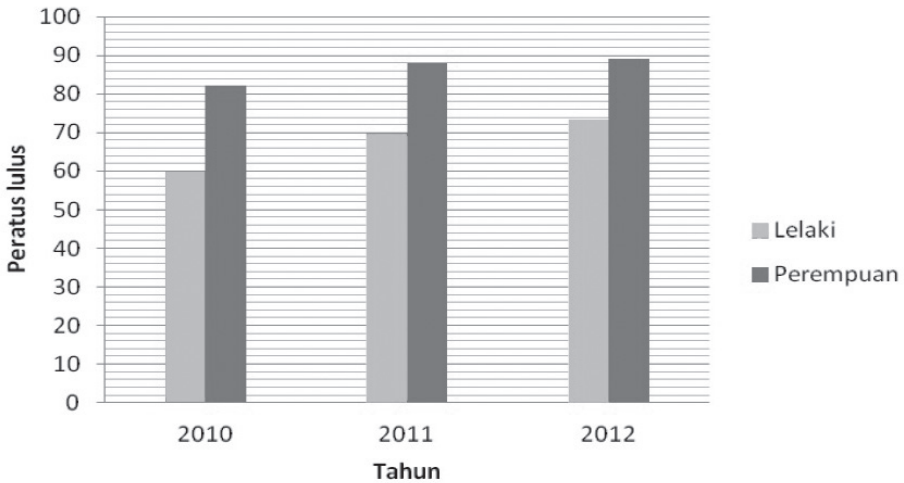
Pernyataan Masalah

Laporan Lembaga Peperiksaan Malaysia (2013) melaporkan bahawa kelulusan peperiksaan SPM mencatatkan peratus kelulusan dari segi lokasi iaitu kawasan bandar dengan luar bandar yang tidak seimbang. Perbezaan yang sangat ketara hasil daripada keputusan SPM mata pelajaran Prinsip Perakaunan bagi negeri Perak dari tahun 2010 hingga 2012 berdasarkan sekolah bandar dan luar bandar (Lembaga Peperiksaan Malaysia, 2013) ditunjukkan dalam Rajah 1. Keputusan sepuluh buah sekolah di bandar dan sepuluh buah sekolah di luar bandar dibandingkan dalam Rajah 1.



Rajah 1: Peratus Kelulusan Mata Pelajaran Prinsip Perakaunan Mengikut Lokasi di Perak. (Lembaga Peperiksaan Malaysia, 2013)

Pada tahun 2010, sekolah di bandar mencatatkan peratus kelulusan sebanyak 89.94% berbanding sekolah di luar bandar dengan peratus kelulusan 63.37%. Pada tahun berikutnya sekolah bandar mencatatkan keputusan yang lebih tinggi (93.36%) berbanding sekolah di luar bandar (77.43%). Pada tahun 2012, sekali lagi sekolah di bandar mendahului sekolah di luar bandar dengan perbezaan peratus kelulusan yang tinggi sebanyak 44.97%. Perbezaan peratus kelulusan yang sangat ketara antara sekolah bandar dengan sekolah luar bandar ini perlu diteliti.



Rajah 2: Peratus Kelulusan Mata Pelajaran Prinsip Perakaunan SPM Mengikut Jantina di Perak. (Lembaga Peperiksaan Malaysia, 2013)

Rajah 2 pula menunjukkan data peratus kelulusan mata pelajaran Prinsip Perakaunan SPM yang berdasarkan jantina yang dikeluarkan oleh Lembaga Peperiksaan Malaysia (2013) untuk negeri Perak. Pada tahun 2010, peratus kelulusan pelajar lelaki adalah 59.9% berbanding dengan peratus kelulusan pelajar perempuan adalah 82.3%. Pada tahun 2011 pula, peratus kelulusan pelajar lelaki adalah 69.9% dan 87.8% bagi pelajar perempuan. Seterusnya pada tahun 2012, peratus kelulusan pelajar lelaki menunjukkan 73.4% dan pelajar perempuan ialah 89.1%. Perbezaan pencapaian yang jelas antara lelaki dengan perempuan dapat dilihat. Kenapakah keadaan ini berlaku berdasarkan bukti analisa hasil keputusan SPM 3 tahun berturut-turut seperti dilaporkan?

Terdapat banyak faktor DIF menyumbang dan mempengaruhi tahap pencapaian pelajar (Nworgu, 2011). Dua faktor utama yang menyumbang kepada perbezaan pencapaian ialah faktor jantina dan lokasi. Doran and Holland (1993) telah mengkaji hubungan pencapaian dalam Ujian Pencapaian Prinsip Perakaunan dengan jantina dan mendapati pencapaian pelajar perempuan dalam Prinsip Perakaunan adalah lebih menonjol. Selain itu, Ogbebor (2012) juga menyatakan wujudnya item yang berfungsi berbeza-beza mengikut lokasi sekolah dalam ujian pencapaian ekonomi (NECO) dan berpihak kepada pelajar bandar.

Terdapat banyak kajian telah dijalankan bagi mata pelajaran sains dan matematik tentang isu DIF tetapi kurang kajian tentang isu ini dalam mata pelajaran Prinsip Perakaunan (Tho, 1994). Selain itu, menurut Boyd, Boyd, & Boyd (2000), kaum wanita lebih mendominasi pekerjaan dalam bidang perakaunan. Daripada masalah-masalah yang dikenalpasti, kajian ini telah mengesan DIF item dalam ujian pencapaian mata pelajaran Prinsip Perakaunan. Antara isu penting yang hendak dikaji ialah adakah item yang digunakan benar-benar berfungsi secara berbeza. Item ujian perlu diuji agar tidak berfungsi secara berbeza-beza terhadap jantina dan lokasi sekolah agar keadilan sosial terjamin.

Objektif Kajian

Kajian ini dijalankan untuk:

- a. Mengesan DIF item-item dalam ujian pencapaian mata pelajaran Prinsip Perakaunan Tingkatan Empat terhadap pelajar-pelajar luar bandar berbanding pelajar-pelajar bandar.
- b. Mengesan DIF item-item dalam ujian pencapaian mata pelajaran Prinsip Perakaunan Tingkatan Empat terhadap pelajar-pelajar lelaki berbanding pelajar-pelajar perempuan.

SOROTAN KAJIAN

Keberbezaan Fungsi Item (*Differential Item Functioning*)

Analisa DIF merujuk kepada kaedah yang lazimnya digunakan untuk menyemak jika sesebuah ujian berat sebelah kepada kumpulan yang diberikan (Camili & Shepard, 1994). Selain itu, DIF merujuk kepada satu set teknik statistik untuk menyemak atau menguji perbezaan prestasi antara kumpulan pelajar terhadap respons item yang diberikan selepas mengawal trait atau ciri-ciri yang diakses (Scherbaum & Goldstein, 2008). Apabila individu daripada kumpulan yang berbeza dengan ciri-ciri yang sama mempunyai kebarangkalian menjawab betul yang sama menghasilkan kesan impak item di mana item tersebut berpihak kepada pelajar dengan kebarangkalian yang tinggi menjawab betul, DIF akan wujud.

DIF yang merujuk kepada bias statistik di mana satu atau lebih parameter daripada modal statistik tersebut disalahtafsirkan atau lebih tafsir (Camili, 2006; Wiberg, 2007). Apabila DIF wujud dalam sesuatu item, punca varians ini perlu dikaji agar tidak berat sebelah (Hossein Karami, 2012). Kehingaran dalam varians (*construct- irrelevant variance*) berpunca daripada faktor kumpulan pelajar diuji. Dalam kajian DIF, terdapat dua kumpulan yang terdiri daripada kumpulan fokal dan kumpulan rujukan. Kumpulan fokal merujuk kepada kumpulan minoriti umpamanya di mana berada dalam kumpulan kurang pencapaiannya. Kumpulan yang berpencapaian lebih baik ialah kumpulan rujukan (Hossein Karami, 2012).

Terdapat dua jenis DIF iaitu DIF seragam dan DIF bukan seragam. DIF seragam wujud apabila satu kumpulan pelajar sentiasa menunjukkan pencapaian yang baik berbanding kumpulan yang lain pada skala kemampuan yang berbeza-beza. DIF bukan seragam pula wujud apabila ahli dalam satu kumpulan menunjukkan kebolehan berbeza pada skala kemampuan yang sama dan wujudnya kesan interaksi antara kumpulan dan skala kemampuan.

Kriteria diperlukan apabila menyesuaikan responden pada skala kebolehan yang biasanya dipanggil '*latent trait*' atau konstruk terpendam. DIF boleh menjadi bukti bias jika dan hanya jika faktor penyebab DIF tersebut tidak

relevan kepada konstruk diukur dalam ujian tersebut. Jika faktor itu sebahagian daripada konstruk, ia dikenali sebagai kesan impak item dan bukannya bias. Panel pakar perlu dirujuk agar kesahan kepada tafsiran lebih jitu.

Ujian Pencapaian Prinsip Perakaunan

Penilaian dalam pencapaian ujian Prinsip Perakaunan perlulah berterusan dan penting dalam memantau proses pembelajaran mata pelajaran ini (Marriot & Lau, 2008). Kajian oleh Tho (1994) telah menyiasat segala faktor-faktor yang menyebabkan perbezaan performa pencapaian dalam peperiksaan Prinsip Perakaunan pelajar di peringkat universiti kursus Perakaunan pengajian tinggi di Universiti Malaya. Menurut Tho (1994) ketiadaan penyelidikan di Malaysia yang diterbitkan dalam Prinsip Perakaunan menjadi perangsang utama kajian tentang ujian pencapaian Prinsip Perakaunan. Analisis regresi berganda menjelaskan 66% daripada variasi dalam pencapaian pelajar disumbangkan oleh gred dalam mata pelajaran matematik dan ekonomi sekolah menengah. Walau bagaimanapun, pembolehubah sosio-demografi jantina dan status kediaman tidak menyumbang kepada kepelbagaian dalam pencapaian pelajar.

Selain itu, kajian Koh dan Koh (2012) menjelaskan impak enam pembolehubah iaitu jantina, pengetahuan sedia ada akaun, aptitud terhadap akademik, latar belakang mata pelajaran matematik, pengalaman bekerja serta umur mempengaruhi pencapaian dalam ujian pencapaian Prinsip Perakaunan pelajar tahun 3 Sarjana Muda Perakaunan. Kajian ini melibatkan 526 sampel dan didapati pencapaian pelajar lelaki lebih menonjol berbanding pelajar perempuan dalam semua aspek kecemerlangan pencapaian akademik dalam subjek Perakaunan. Selain itu, pengaruh teori budaya perlu dipertimbangkan dalam mengukur pencapaian dalam Prinsip Perakaunan supaya pengukuran prestasi lebih baik (Gray, 2005).

Gallagher (2001) menyatakan perbezaan pencapaian di antara jantina disebabkan oleh interaksi faktor budaya dan sosial. Kajian oleh Jackling dan Anderson (1998), menjelaskan bahawa pencapaian pelajar perakaunan pelajar separuh masa lebih baik berbanding pelajar sepenuh masa. Dima dan Mbekomize (2013) pula mendapati pelajar perempuan lebih tinggi pencapaian berbanding pelajar lelaki kerana faktor sikap dan komitmen tinggi terhadap pencapaian Prinsip Perakaunan. Okafor dan Egbon (2011) menyatakan tiada perbezaan dalam pencapaian ujian Perakaunan di Nigeria. Ross (2009) bersependapat dengan Dima dan Mbekomize (2013) bahawa pelajar perempuan lebih tinggi pencapaian dalam ujian pencapaian Prinsip Perakaunan (CIMA) di Afrika. Brahmasrene dan Whitten (2001) pula melaporkan pelajar lelaki lebih tinggi pencapaian dalam ujian pencapaian Prinsip Perakaunan berbanding pelajar perempuan. Kajian longitunal oleh Sheard (2009) pula mendapati pencapaian pelajar perempuan adalah lebih baik berbanding dengan pelajar lelaki dalam perakaunan di peringkat pengajian tinggi disebabkan komitmen dan sikap terhadap akademik.

Kerangka Teori

Teori yang digunakan untuk kajian ini ialah Teori Gerak Balas Item (*Item Response Theory - IRT*). Teori ini memberi penekanan kepada tingkah-laku individu apabila memberikan respons kepada item-item ujian. Konsep asas bagi pendekatan Teori Gerak Balas Item bergantung kepada item-item secara individu berbanding teori klasik yang bergantung kepada aggerat bagi respon-respon untuk item seperti markah ujian (Baker, 2001). Model Rasch satu parameter diaplikasikan untuk tujuan kajian ini yang mengandaikan semua item adalah mempunyai darjah diskriminasi yang sama serta item tidak dapat dijawab melalui tekaan (Angoff, 1993). Model Rasch ialah model matematik berdasarkan teori kebarangkalian. Keluk Ciri Item menyediakan kebarangkalian pelajar menjawab betul pada titik yang berbeza pada skala kecekapan yang sama memudahkan perbandingan di antara sub kumpulan dilakukan.

Fokus utama Teori Respons Item dalam mengesan DIF ialah parameter item dan Keluk Ciri Item (*Item Characteristic Curve - ICC*) seperti dicadangkan oleh Camilli dan Shepard (1994). Kaedah Teori Gerak Balas Item dalam mengesan DIF membenarkan pengesanan item-item ujian yang tidak berfungsi serupa bagi kumpulan pelajar yang berlainan melalui perbandingan Keluk Ciri Item. DIF disyaki wujud apabila Keluk Ciri Item bagi item yang sama berbeza-beza untuk dua kumpulan yang dibandingkan melebihi ralat persampelan. Sebelum memulakan analisis pengesanan DIF menggunakan Model Rasch, andaian dibuat bahawa semua item dalam ujian mematuhi kehendak model seperti unidimensional, berhierarchy, linear dan objektif. Nilai kebolehppercayaan pengasingan item (*Item separation reliability*) mencapai nilai lebih daripada 0.90 dapat memenuhi ciri unidimensional. Item-item yang mempunyai nilai 'mean square' (MNSQ) bagi infit dan outfit mestilah dalam julat antara 0.7 hingga 1.3 logits. Item-item yang berada di luar julat ini berkemungkinan mengukur kebingaran yang tidak berkaitan dengan ujian. Keserasian dengan model diperolehi dengan nilai keserasian *chi-square* berdasarkan frekuensi yang diperhatikan atau dijangkakan untuk skor diperolehi (Orlando & Thissen, 2000).

Selepas menyemak keperluan model iaitu unidimensional dan keserasian dengan model, prosedur 'anchoring' atau meletakkan nilai pautan dilaksanakan. Prosedur ini penting untuk mengesan DIF bagi setiap item agar menghasilkan piawaian yang tinggi untuk perbandingan min antara kumpulan. Dengan ini, item-item tersebut akan mempunyai nilai min menghampiri 0 dan SD menghampiri 1 untuk dijadikan taburan normal iaitu $t(0,1)$. Satu translasi penting dalam menghasilkan satu asalan yang biasa kepada kedua-dua set item bagi menganggarkan kesukaran yang sama bagi perbandingan antara kumpulan.

Item-item peperiksaan Prinsip Perakaunan SPM menepati kehendak model agar sesuatu ujian atau pembolehubah adalah terbina daripada item-item yang mempunyai persamaan dari segi konstruk yang hendak diukur. Proses mengesan item-item yang berfungsi berbeza-beza bagi kumpulan berdasarkan

lokasi (bandar dan luar bandar) serta jantina (lelaki dan perempuan) adalah berdasarkan nilai pautan yang telah diletakkan sebelumnya. Kandungan ujian perlulah dianalisa dengan teliti berdasarkan sukatan pelajaran Prinsip Perakaunan terkini supaya tidak mempengaruhi pengesanan DIF (Lane, Wang, & Magone, 1996). Item yang berbentuk aneka pilihan lazimnya menggunakan jumlah skor diperolehi dalam ujian untuk menganggarkan skala kebolehan pelajar. Berdasarkan nilai $t \geq \pm 1.96$ kewujudan DIF dapat dikesan secara statistik.

Distribusi hipotetikal untuk kumpulan perbandingan diberikan pada skala kebolehan yang sama bagi setiap item memberikan bukti DIF bagi item tersebut. Perbezaan dalam pencapaian item diukur dengan membandingkan kebarangkalian menjawab betul merentasi pembolehubah dikehendaki dalam subpopulasi (Penfield & Lam, 2000). Skor yang diperolehi melambangkan pencapaian dalam Prinsip Perakaunan boleh digunakan sebagai pembolehubah yang dikehendaki dalam kajian. Perbezaan pencapaian yang dilihat pada parameter kesukaran item dalam Model Rasch diterjemahkan dalam bentuk gambarajah dengan memplotkan 'scatter plot' bagi sub-kumpulan lokasi dan jantina terlibat. Item-item yang mempunyai parameter kesukaran item melebihi ± 0.5 logit bagi setiap subkumpulan diperhatikan dan saiz DIF dapat dikesan. Dalam ujian pendidikan untuk pencapaian, item-item yang berbeza skala dalam sub kumpulan adalah dikatakan berfungsi berbeza-beza atau DIF kerana item didapati agak sukar bagi satu sub kumpulan berbanding sub kumpulan yang lain (Andrich, 1988). Aspek pembangunan item ujian pencapaian perlulah mengambil kira faktor kebebasan konstruk item, pentadbiran ujian yang teratur dan rubrik penskoran (Yen, 1993) agar kebebasan setempat lebih mantap.

METODOLOGI

Sebanyak 15 buah sekolah di luar bandar dan 5 buah sekolah di bandar sekitar Ipoh terlibat dengan kajian ini. Sampel yang terlibat dalam kajian terdiri daripada pelajar Tingkatan Empat yang mengambil Prinsip Perakaunan di sekolah-sekolah negeri Perak. Seramai 1000 sampel telah terlibat dalam kajian ini. Sampel-sampel yang dipilih terdiri daripada pelbagai kebolehan dan mengikuti pembelajaran mata pelajaran Prinsip Perakaunan yang sama. Taburan sampel mengikut jantina dan lokasi sekolah adalah seperti dalam Jadual 1.

Jadual 1: Taburan Sampel dalam Kajian

Demografi	Kategori	Frekuensi	Peratus(%)
Jantina	Lelaki	487	48.7
	Perempuan	513	51.3
Jenis sekolah	Luar Bandar	482	48.2
	Bandar	518	51.8

Pengkaji telah membina satu ujian pencapaian yang mengandungi 40 item pelbagai pilihan yang dipilih secara rawak daripada koleksi item kertas 1 Prinsip Perakaunan SPM daripada tahun 2005-2010. Menurut Bahagian Pembangunan Kurikulum (2015) ujian pencapaian kertas 1 Prinsip Perakaunan mengandungi aras kognitif pengetahuan, pemahaman dan aplikasi yang sesuai dengan pelajar Tingkatan Empat dan item-item disusun mengikut aras kesukaran bernisbah 6:3:1 (senang: serderhana:susah) (Lembaga Peperiksaan Malaysia, 2015). Tiga orang guru cemerlang dan pakar mata pelajaran Prinsip Perakaunan serta mempunyai pengalaman mengajar melebihi minima 10 tahun dirujuk untuk mendapatkan pengesahan aras kesukaran ini. Item-item dalam soalan Prinsip Perakaunan disusun dalam bentuk pendekatan pernyataan, diagramatik, pengiraan, situasi, penyelesaian masalah dan sebagainya berdasarkan Kurikulum Standard Prinsip Perakaunan. Ini kerana pelbagai bentuk item yang diutarakan sebagai memberi rangsangan minda kepada calon dan ia adalah mirip soalan peperiksaan SPM sebenar. Ia juga bertujuan untuk mendapatkan kesahan kandungan bagi ujian tersebut.

Item-item ini ditadbirkan kepada sampel dan dijawab dalam masa 1 jam 15 minit. Arahan dan prosedur pentadbiran ujian disediakan untuk memastikan ujian-ujian ditadbir dengan standard. Program analisis komputer Model Rasch: WINSTEPS digunakan untuk menganalisis data. Terlebih dahulu program ini mesti menentukan sama ada data ujian padan dengan Model Rasch (Andrich, 2004). Hasil ujian telah dianalisis menggunakan Model Pengukuran Rasch dan perisian WINSTEP serta laporan hasil kajian dilaporkan.

DAPATAN KAJIAN

Jadual 2 menunjukkan analisa DIF merentasi sekolah di bandar dan sekolah di luar bandar. Terdapat item-item yang dikesan mempunyai DIF bagi kumpulan pelajar di luar bandar dan bandar. Sebanyak 17 item didapati berfungsi secara berbeza mengikut lokasi sekolah. Item-item suspek DIF yang bertindak berbeza-beza ini ialah item 2, 3, 8, 9, 11, 13, 14, 16, 17, 19, 25, 26, 27, 28, 31, 32, dan 39. Item-item ini dilabelkan sebagai DIF kerana nilai $t \geq \pm 1.96$ pada aras $p = .05$. Linacre (2011) menyatakan bahawa item-item yang $t \geq \pm 1.96$ pada aras $p = .05$ merupakan item suspek DIF.

Jadual 2: Keberbezaan Fungsi Item antara Lokasi

Kumpulan (Bandar)	DIF Measure (Kesukaran item)	Kumpulan (Luar Bandar)	DIF Measure (Kesukaran item)	DIF Contrast (Saiz DIF)	Nilai t	No Item
B	0	L	-1	0	*3.66	2
B	0	L	0	0	*2.87	3

Kumpulan (Bandar)	DIF Measure (Kesukaran item)	Kumpulan (Luar Bandar)	DIF Measure (Kesukaran item)	DIF Contrast (Saiz DIF)	Nilai t	No Item
B	0	L	0	0	*3.68	8
B	1	L	0	0	*4.30	9
B	0	L	0	0	*3.59	11
B	0	L	0	0	*-2.63	13
B	0	L	0	0	*2.94	14
B	0	L	0	0	*-2.14	16
B	0	L	0	0	*3.57	17
B	0	L	0	0	*-2.68	19
B	0	L	0	0	*-3.30	25
B	0	L	0	0	*-2.20	26
B	0	L	0	0	*-3.88	27
B	0	L	0	0	*-3.50	28
B	0	L	0	0	*-4.15	31
B	0	L	0	0	*-2.33	32
B	0	L	1	0	*-2.84	39

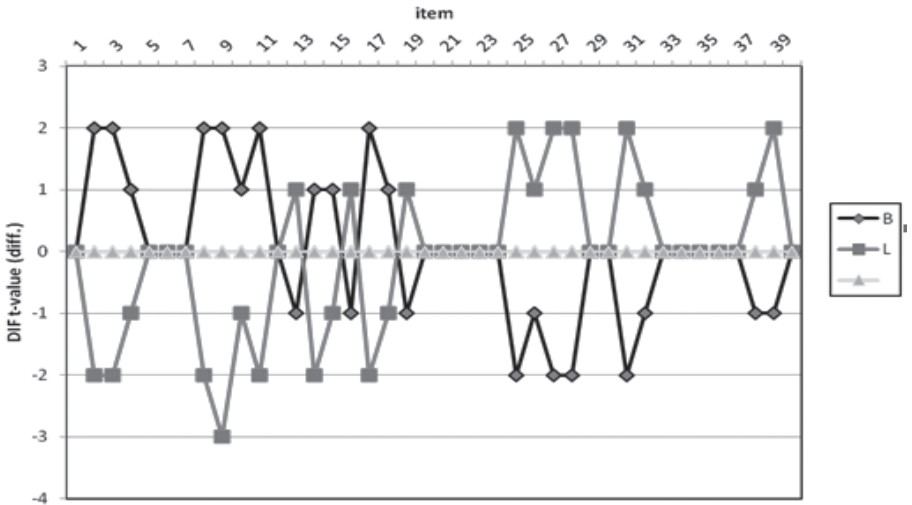
* Item DIF berdasarkan nilai $\geq \pm 1.96$ pada $p = .05$

Jadual 3 menunjukkan item-item DIF mengikut lokasi sekolah bandar dan luar bandar. DIF measure merujuk kepada perbezaan signifikan antara dua nilai indeks kesukaran bagi dua kelompok atau kumpulan. Manakala DIF contrast merujuk kepada perbezaan antara indeks kesukaran item bagi setiap kelompok atau kumpulan. Ia adalah penting untuk mengesan peringkat pengesahan jurang perbezaan indeks kesukaran setiap item. *DIF measures* mengesahkan item berdasarkan anggaran kelompok pelajar sekolah luar bandar dan bandar. Lebih sukar sesuatu item untuk disahkan, lebih besar saiz *DIF Measure* (Linacre, 2011). Dalam kajian ini, *DIF contrast* semua item adalah sifar menunjukkan DIF yang dikesan boleh diabaikan. *DIF contrast* ialah ukuran perbezaan dalam DIF measure dua kelompok pelajar. Menurut Bond dan Fox (2007), DIF contrast yang berada di antara julat $\pm .5$ di antara kelompok pelajar merupakan item yang boleh diabaikan. Berdasarkan Jadual 2 didapati item 2 and item 39 lebih susah untuk pelajar luar bandar berbanding pelajar bandar. Item 9 lebih susah untuk pelajar bandar berbanding pelajar luar bandar.

Jadual 3: Item-item DIF mengikut Lokasi Sekolah Bandar dan Luar Bandar

Kelompok Pelajar	Item DIF	%
Luar Bandar	2,39	5
Bandar	9	2.5

Rajah 3 menunjukkan plot kedudukan item yang dikesan sebagai DIF berdasarkan nilai $t \geq \pm 1.96$. Menurut Bond dan Fox (2007), DIF *contrast* yang melebihi $\pm .5$ di antara kelompok pelajar dibandingkan boleh dikatakan item yang sangat DIF. Secara kesimpulannya, sebanyak 17 item dikesan sebagai item DIF secara statistik tetapi berdasarkan nilai magnitud DIF *contrast* boleh dikatakan tiada item yang menyumbang kepada DIF secara membina.



Rajah 3: DIF lokasi

Jadual 4: Keberbezaan Fungsi Item antara Jantina

Kumpulan (Lelaki)	DIF Measure (Kesukaran item)	Kumpulan (Perempuan)	DIF Measure (Kesukaran item)	DIF Contrast (Saiz DIF)	Nilai t	No Item
L	0	P	0	0	*2.75	8
L	0	P	-1	0	*2.12	10
L	0	P	0	0	*1.96	12
L	0	P	0	0	*-2.71	17
L	1	P	1	0	*-2.13	20
L	0	P	0	0	*3.56	33
L	0	P	0	0	*-2.83	38
L	0	P	0	0	*-2.81	39

* Item DIF berdasarkan nilai $t \geq \pm 1.96$ pada $p=0.05$

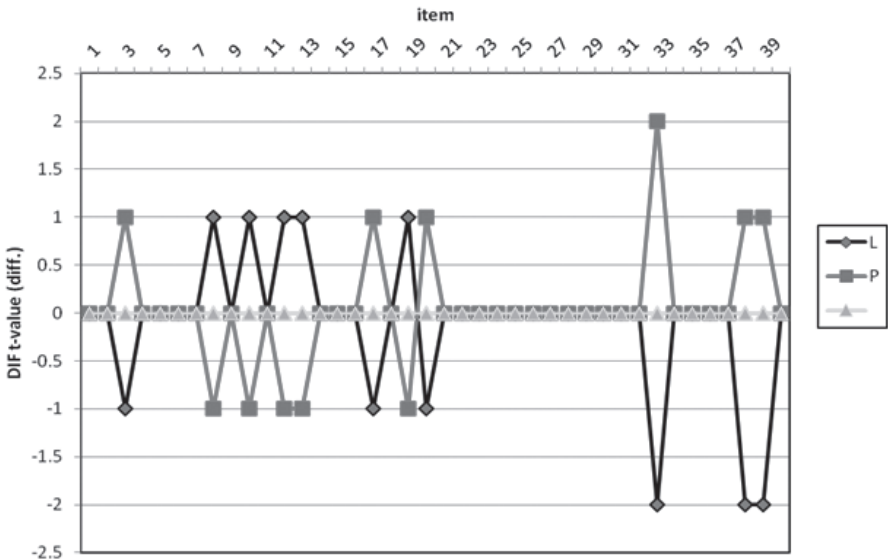
Analisis diteruskan dengan melihat keberbezaan fungsi item, DIF antara pelajar-pelajar lelaki dengan pelajar-pelajar perempuan untuk menjawab soalan kajian seterusnya. Item-item yang didapati berfungsi berbeza-beza kepada kelompok pelajar lelaki dan perempuan ialah item 8, 10, 12, 17, 20, 33, 38 dan

39 berdasarkan Jadual 4. Item ini secara statistikalnya adalah item suspek DIF berdasarkan nilai $t \geq \pm 1.96$ pada aras $p=.05$. DIF *contrast* (saiz DIF) ialah sifar. Keputusan ini jelas menunjukkan DIF yang dikesan boleh diabaikan dan ujian pencapaian tersebut tidak memihak kepada lelaki atau perempuan.

Jadual 5: Item-item DIF mengikut Jantina

Kelompok Pelajar	Item DIF	%
Lelaki	-	0
Perempuan	10	2.5

Berdasarkan Jadual 5, didapati tiada item yang memihak kepada kelompok pelajar lelaki. Item 10 lebih susah untuk kelompok pelajar perempuan berbanding pelajar lelaki. Saiz DIF yang dikesan boleh diabaikan. Rajah 4 menunjukkan plot kedudukan item yang dikesan sebagai DIF berdasarkan nilai $t \geq 1.96$.



Rajah 4: DIF jantina

PERBINCANGAN DAN KESIMPULAN

Daripada analisis yang dijalankan didapati sebanyak 17 item suspek dalam ujian pencapaian ini berfungsi secara berbeza-beza mengikut faktor lokasi sekolah. Item-item 2, 3, 8, 9, 11, 13, 14, 16, 17, 19, 25, 26, 27, 28, 31, 32 dan 39 didapati berfungsi secara berbeza-beza dan menunjukkan keberbezaan fungsi item pada aras $p=.05$ dengan $t \geq \pm 1.96$. Walau bagaimanapun, berdasarkan nilai DIF *contrast* (± 0.5) menunjukkan saiz DIF yang boleh diabaikan di antara

kelompok pelajar bandar dan luar bandar. Didapati item 2 dan item 39 lebih sukar untuk pelajar luar bandar manakala item 9 lebih sukar kepada pelajar bandar berdasarkan nilai DIF measure yang dilaporkan.

Dapatan kajian ini selari dengan dapatan kajian oleh Ogbebor (2012) yang melaporkan bahawa terdapatnya item DIF yang dikesan secara statistik berdasarkan lokasi dalam Peperiksaan Awam tetapi saiz DIF yang dikesan boleh diabaikan kerana magnitudnya kecil. Menurut Roever (2005), mengenal pasti item yang berfungsi berbeza-beza mengikut kelompok pelajar merupakan langkah pertama mengesan item yang berat sebelah. Tambahan pula, kajian oleh Obe (1984), menunjukkan pencapaian pelajar di sekolah bandar lebih menonjol daripada sekolah di luar bandar kerana item lebih berpihak kepada mereka.

Daripada analisis yang dijalankan didapati sebanyak lapan item suspek DIF dalam ujian pencapaian ini berfungsi secara berbeza-beza mengikut faktor jantina. Item-item 8, 10, 12, 17, 20, 33, 38 dan 39 didapati berfungsi secara berbeza-beza dan menunjukkan keberbezaan fungsi item pada aras $p=0.05$ dengan $t \geq \pm 1.96$. Walaubagaimanapun, berdasarkan nilai DIF *contrast* (± 0.5) menunjukkan saiz DIF yang dikesan boleh diabaikan antara kelompok pelajar lelaki dan perempuan. Didapati item 10 lebih sukar untuk pelajar perempuan berbanding pelajar lelaki.

Dapatan soalan kajian ini selari dengan dengan dapatan kajian oleh Gierl, Khaliq, dan Boughton (1999) yang melaporkan tiada magnitud DIF yang besar yang dikesan berfungsi berbeza-beza dalam ujian aneka pilihan yang ditadbirnya. Ini bermaksud skor pelajar dalam ujian yang ditadbir ditentukan oleh kebolehan pelajar yang sama merentasi kumpulan jantina. Dapatan kajian ini selari juga selari dengan dapatan kajian oleh Parker (2005) yang melaporkan tiada magnitud saiz DIF yang besar dikesan dalam ujian pencapaian yang ditadbirnya di antara lelaki dan perempuan di negara Afrika dalam ujian pencapaian ekonomi.

Implikasi daripada kajian ini menyumbang kepada pengukuran pendidikan dalam mata pelajaran Prinsip Perakaunan. Dapatan kajian memberikan refleksi bahawa alat ukur yang digunakan menyumbang secara membina. Pencapaian dalam Prinsip Perakaunan perlu diberikan perhatian agar masa depan mata pelajaran ini lebih baik. Kajian ini akan menjadi panduan dan mendorong guru-guru semasa membuat kesimpulan berhubung item yang perlu diadakan dalam ujian pencapaian. Pemilihan item perlulah bersesuaian dengan segenap lapisan pelajar dalam pengujian. Item-item yang perlu diubahsuai atau disingkirkan sama sekali perlu dipertimbangkan agar penilaian melibatkan seluruh Malaysia adalah adil untuk semua pelajar.

Punca-punca item dalam ujian pencapaian yang memihak kepada pelajar lelaki atau perempuan perlu dikaji dan diperbetulkan. Item-item yang dikesan ini akan membantu penggubal item membina garis panduan agar item tersebut tidak diadakan dalam ujian pencapaian. Garis panduan bagi pembinaan item akan menghasilkan penilaian yang lebih adil dalam mengukur pencapaian

pelajar dalam Prinsip Perakaunan. Ujian pencapaian yang lebih bersifat objektif dapat dihasilkan.

Punca-punca item dalam ujian pencapaian yang memihak kepada pelajar sekolah luar bandar atau pelajar sekolah bandar perlu diperhalusi agar memotivasikan kedua-dua kelompok pelajar mengambil matapelajaran Prinsip Perakaunan. Tambahan pula, dalam kajian ini didapati banyak item yang secara statistik adalah item suspek DIF kepada kelompok pelajar berdasarkan lokasi. Proses membina ujian pencapaian oleh guru dalam mengukur pencapaian pelajar dalam Prinsip Perakaunan perlu diberikan perhatian agar ujian yang dibina benar-benar mengukur pencapaian pelajar. Guru-guru perlu lebih teliti dalam memilih item yang berkualiti dalam ujian pencapaian.

RUJUKAN

- Andrich, D. (1988). *Rasch models for measurement*. Sage Publications, INC.
- Andrich, D. (2004). Controversy and the Rasch model: a characteristic of incompatible paradigms? *Medical care*, 42 (1), 1-7.
- Angoff, W. H. (1993). *Perspectives on differential item functioning methodology*. In P. W. Holland & H. Wainer (Eds.), *Differential item functioning* (pp. 3-4). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Bahagian Pembangunan Kurikulum (2015). *Huraian Sukatan Pelajaran Tingkatan 4 dan 5*. Retrieved March 27, 2015, from <http://apps2.moe.gov.my/kurikulum/v2/index.php>
- Baker, F. (2001). *The Basics of Item Response Theory*. ERIC Clearinghouse on Assessment and Evaluation, University of Maryland, College Park, MD.
- Barrett, D. W., (2001, August). Religious orientation and social cognitive motivation. Paper presented at the *109th American Psychological Association Annual Convention*, San Francisco, CA.
- Bond, T. G. dan Fox, C. M. (2001). *Applying the Rasch model: Fundamental measurement in the human sciences*. Mahwah, NJ: Erlbaum.
- Bond, T.G. DAN Fox, C. M. (2007). *Applying the Rasch Model: Fundamental Measurement in the Human sciences*. New Jersey: Lawrence Erlbaum Associates.
- Boyd, D. T., Boyd, S.C. dan Boyd, W. L. (2000, September/ October). Changes in accounting education: Improving principles content for better understanding. *Journal of Education for Business*, 76(1), 1-2. Retrieved November 17, 2003, from Business Source Premier database (3871722).
- Brahmasrene, T., dan Whitten, D. (2001). Assessing success on the uniform CPA exam: A logit approach. *Journal of Education for Business*, 77(1), 45-50.
- Brown, F. G. (1983). *Principles of educational and psychological testing* (3rd ed.). New York: Holt, Rinehart, and Winston.
- Camilli, G. dan Shepard, L.A. (1994). *Methods for identifying biased test items*. Jaeger RM, editor. Thousand Oaks: Sage.

- Camilli, G. (2006). Test fairness. In R. Brennan (Ed.). *Educational measurement* (pp. 221-256).
- Dima, L.W. dan Mbekomize, C.J.(2013). Causes of Gender Differences in Accounting Performance: Students' Perspective. *International Education Studies*, 6, 13-26.
- Dimitrov, D.M. (1999). Gender differences in science achievement: Differential effect of ability, response format, and strands of learning outcomes. *School Science & Mathematics*, 99(8), 445-450.
- Dorans, N. J., dan Holland, P.W. (1993). DIF detection and description: Mantel-Haenszel and Standardization. In P. W. Holland & H. Wainer (Eds.), *Differential item functioning* (pp. 35-66). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Ercikan, K. (2002). Disentangling sources of differential item functioning in multilanguage assessments. *International Journal of Testing*, 2(3&4), 199-215.
- Gallagher, T. (2001). *Equal opportunities commission conference on boys and girls in the 21st Century: Gender Differences in Learning*. Retrieved February 2, 2013, from <http://www.eoc.org.hk/eoc/graphicsfolder/inforcenter/newsletter/content.aspx?itemid=4948>
- Gray, S. J. (2005). Towards a Theory of Cultural Influence on the Development of Accounting Systems Internationally. *A Journal of Accounting, Finance and Bussiness Studies*, 24(1), 1-15.
- Grierl, M., Khaliq, S. H., dan Boughton, K. (1999). Gender differential item functioning in mathematics and science: prevalence and policy implications. Paper presented at the *Symposium Improving large-scale assessment in education* at the annual meeting of the Canadian society for the Study of Education.
- Hanushek, E. A. dan Kimko, D. D. (2000). Schooling, labor-force quality, and the growth of nations. *American Economic Review*, 90(5), 1184-1208.
- Hossein Karami (2012). An introduction to Differential Item Functioning. *The International Journal of Educational and Psychological Assessment*, 11(2), 59-76.
- Jackling, B., and Anderson, A. (1998). Study mode, general ability and performance in accounting: A research note, *Accounting Education* 7, 65-73.
- Koh, M. Y. dan Koh, H. C.(2012). The determinants of performance in an accountancy degree programme. *Accounting Education*, 8 (1), 13- 29.
- Lane, S., Wang, N., & Magone, M. (1996). Gender related differential item functioning on a middle school mathematics performance assessment. *Educational Measurement: Issues and Practice*, 15(4), 21-27, 31.
- Lembaga Peperiksaan Malaysia (2013). *Peperiksaan dan Pentaksiran*. Retrieved March 27, 2013, from <http://apps2.moe.gov.my/lponline/v1/index.php?lang=en>
- Linacre, J. M. (2008). *A user's guide to WINSTEPS-MINISTEP: Rasch-model computer programs*. Chicago, IL: winsteps.com.

- Linacre, J. M. (2010). Winsteps® (Version 3.70.0) [Computer Software]. Beaverton, Oregon:Winsteps.com.
- Linacre, J. M. (2011). Winsteps Rasch Measurement Version 3.71 [Software]. Available from:<http://www.winsteps.com>.
- Maksy, M. M. & Lin Zheng (2008). Factors associated with student performance in advanced accounting and auditing - an empirical study in a public university. *Accounting Research Journal*, 21(1), 16-32.
- Linacre, J. M. (2013). Differential Item Functioning DIF Sample Size Nomogram, *Rasch Measurement Transactions*, 26:4 p. 1391.
- Marriott, P. & Lau, A. (2008). The use of on-line summative assessment in an undergraduate financial accounting course. *Journal of Accounting Education*, 26,73-90.
- Maksy, M. M. & Lin Zheng (2008). Factors associated with student performance in advanced accounting and auditing - an empirical study in a public university. *Accounting Research Journal*, 21(1), 16-32.
- Nworgu, B. G. (2011). Differential item functioning :A critical issue in regional quality assurance. Paper presented in NAERA conference.
- Obe, E.O. (1984). Urban-rural and sex difference in scholarstic aptitude on primary school finalist in Lagos state. *Education and Development*, 41(2), 123-134
- Obinne, A.D.E., Nworgu, B.G.,& Umobong, M.E. (2013). An Investigation into Differential item Functioning of Tests Conducted by the Two major Examination Bodies in Nigeria. *Advances in Educational Research*, 2 (1),001-008.
- Ogbebor, U. C. (2012). *Differential Item Functioning Economics Question Paper of National Examinations*, Council in Delta State Nigeria.
- Okafor, C. A., & Egbon, O. (2011). Academic Performance of male versus Female Accounting Undergraduate Students: Evidence from Nigeria. *Higher Education Studies*, 1(1), 9-19.
- Orlando, M., & Thissen, D. (2000). New item fit indices for dichotomous item response theory models. *Applied Psychological Measurement*, 24, 50-64.
- Parker, G. (2005). Beyond major depression. *Psychol Med*. 35, 467– 474.
- Penfield, R. D., & Lam, T. C. M. (2000). Assessing differential item functioning in performance assessment. *Educational Measurement: Issues and Practice*, 19(3), 5–15.
- Roeber, C. (2005). “That’s not fair!” *Fairness, bias, and differential item functioning in language testing*. Retrieved November 18, 2013, from the University of Hawai’i System Web site: <http://www2.hawaii.edu/~roeber/brownbag.pdf>.
- Ross, S. (2009). Factors affecting Southern African students’ success in CIMA examinations. *Meditari Accountancy Research*, 17(1), 49-67.
- Scherman, C. A., & Goldstein, H. W. (2008). Examining the relationship between race-based Differential Item Functioning and Item Difficulty. *Educational and Psychological Measurement*; 68, 537-553.

- Sheard, M. (2009). Hardiness, commitment, gender, and age differentiate university academic performance. *British Journal of Educational Psychology*, 79, 189-201.
- Siti Rahayah Ariffin. (2003). *Teori, Konsep dan Amalan Dalam Pengukuran dan Penilaian*. Pusat Pembangunan Akademik Bangi: Penerbit Universiti Kebangsaan Malaysia.
- Tho, L.M. (1994). Some evidence on the determinants of student performance in the University of Malaya introductory accounting course'. *Accounting Education*, 3(4), 331-340.
- Wiberg, M. (2007). Measuring and detecting differential item functioning in criterion-referenced licensing test. A Theoretic Comparison of Methods. *Educational Measurement*, technical report N. 2.
- Yen, W.M. (1993). Scaling performance assessments: Strategies for managing local item dependence. *Journal of Educational Measurement*, 30, 187-213.
- Alpkaya, U., & Koceja, D. (2007). The effects of Acute Static Stretching On reaction Time & Force. *Journal of Sports Medicine & Physical Fitness*; 47, 2.