



**PENDEKATAN MODEL ITEM RESPONSE THEORY (IRT)
DALAM PEPERIKSAAN MENGGUNAKAN KOMPUTER
DALAM KALANGAN PENSYARAH UTHM**

**MOHAMAD ZAID MUSTAFA
ALI SURADIN
FAIZAH HANI ABDUL RASHID
KAHIROL MOHD SALLEH
NOR LISA SULAIMAN
BADARUDDIN IBRAHIM
MOHD SAFIEE IDRIS**

PENDEKATAN MODEL *ITEM RESPONSE THEORY* (IRT) DALAM PEPERIKSAAN MENGGUNAKAN KOMPUTER DALAM KALANGAN PENSYARAH UTHM

Mohamad Zaid Mustafa (mzaid@uthm.edu.my), Ali bin Suradin, Faizah Hani Abdul Rashid, Kahirol Mohd. Salleh (kahirol@uthm.edu.my), Nor Lisa Sulaiman, Badaruddin Ibrahim (badar@uthm.edu.my) dan Mohd. Safiee Idris
Fakulti Pendidikan Teknikal, Universiti Tun Hussein Onn Malaysia, 86400 Parit Raja, Batu Pahat, Johor

Abstrak

Kajian ini bertujuan untuk meninjau persepsi pensyarah Universiti Tun Hussein Onn Malaysia (UTHM) terhadap pendekatan Model IRT (*Item Response Theory*) dalam peperiksaan menggunakan komputer. Kajian turut mengenalpasti tahap kefahaman pensyarah UTHM terhadap penggunaan komputer dalam peperiksaan dengan pendekatan model IRT, persepsi pensyarah terhadap sistem peperiksaan yang mengenalpasti kemahiran dan kebolehan pelajar serta kebolehlaksanaannya diaplikasikan di dalam sistem pendidikan negara kita. Kajian dijalankan melalui soal selidik dan melibatkan responden yang terdiri daripada 86 orang pensyarah di tiga fakulti iaitu Fakulti Pendidikan Teknikal, Fakulti Pengurusan Teknologi dan Fakulti Teknologi Maklumat & Multimedia, UTHM. Keseluruhan data yang diperolehi dianalisis menggunakan perisian *Statistical Package For Social Sciences (SPSS)* ver 11.0 dan analisa dibuat dalam bentuk peratusan, skor min dan sisihan piawai. Dapatan kajian menunjukkan tahap kefahaman pensyarah terhadap penggunaan komputer dalam peperiksaan adalah tinggi. Namun demikian, pensyarah menunjukkan persepsi yang sederhana terhadap pendekatan model IRT dalam peperiksaan dan ciri-ciri model tersebut sebagai kaedah mengenalpasti kemahiran pelajar. Kebolehlaksanaan model IRT dalam peperiksaan berada ditahap yang rendah menunjukkan bahawa pensyarah kurang bersetuju supaya sistem tersebut diaplikasikan. Oleh itu, beberapa cadangan telah dikemukakan untuk kajian akan datang dalam memberi pendedahan berkaitan dengan model IRT.

PENGENALAN

Sistem peperiksaan di Malaysia secara keseluruhannya menggunakan kaedah CTT (*Classical Test Theory*), di mana peperiksaan ini menggunakan kaedah tradisi iaitu dengan menggunakan kertas dan pensil. Menurut Wiberg (2004), teori peperiksaan secara tradisi telah menguasai sistem peperiksaan yang telah dipiawaikan sejak sekian lama dengan berdasarkan kaedah pemarkahan yang perlu diteliti oleh setiap pelajar. Di institusi pendidikan, peperiksaan yang dijalankan terhadap pelajar adalah untuk menentukan kedudukan seseorang pelajar itu dimana tahap prestasi pelajar itu dinilai dan dibandingkan dengan prestasi pelajar yang lain berdasarkan keputusan daripada peperiksaan tersebut (Tom & Gary, 2003). Kaedah ini dipanggil *Norm-Referenced Test* (NRT). Kaedah ini menjadikan guru hanya berupaya menyimpan data dan maklumat peperiksaan pelajar sahaja sebagai rujukan.

Kebolehan sistem peperiksaan mengukur tahap kecekapan pelajar dalam menguasai sesuatu kemahiran merupakan salah satu teknik dalam membantu pelajar membuat keputusan samada memerlukan usaha yang lebih atau kurang bergantung kepada keupayaan pelajar tersebut (Tom & Gary, 2003). Kaedah ini dipanggil *Criterion-Referenced Test* (CRT), di mana peperiksaan yang dijalankan adalah sebagai satu alat untuk mengenalpasti kebolehan dan kemahiran yang terdapat pada pelajar. Keupayaan dan kemahiran pelajar terhadap sesuatu subjek akan dinilai dan diuji dengan soalan yang setaraf. Bermula dari soalan yang mudah, sederhana dan sukar untuk dijawab mengikut kepada kebolehan seseorang pelajar. Ciri-ciri berikut adalah merupakan salah satu karakter di dalam model IRT iaitu *Item Response Theory*.

Item Response Theory (IRT) adalah merupakan kerangka statistik di mana pelajar yang mengambil peperiksaan boleh digambarkan samada mempunyai satu kemahiran atau lebih melalui model-model matematik yang menghubungkan prestasi sebenar pada item peperiksaan, item statistik dan kebolehan pelajar. Menurut Patricia (1999), model IRT diasaskan oleh Frederick Lord bermula pada awal 1970an lagi. Kejayaan inilah yang menghasilkan suatu penemuan tentang model yang berdasarkan kaedah matematik untuk kaedah pengukuran yang *standard* di dalam pendidikan. Kekuatan IRT adalah bergantung kepada ciri-ciri parameter yang tidak pelbagai dan boleh diaplikasikan di dalam program komputer yang *standard*. IRT mempunyai tempat di mana soalan-soalan peperiksaan dibina, dikumpul dan ditempatkan iaitu dipanggil "*item banking*". Ia merupakan koleksi yang terbesar dimana setiap guru atau pembuat soalan boleh menghasilkan set-set soalan yang berpadanan dengan keperluan pengukuran yang spesifik. Penyimpanan maklumat peperiksaan di dalam "*item banking*" adalah merupakan pangkalan data elektronik dimana terdapat pembinaan, pengurusan, pemilihan, dan penapisan soalan-soalan yang disimpan.

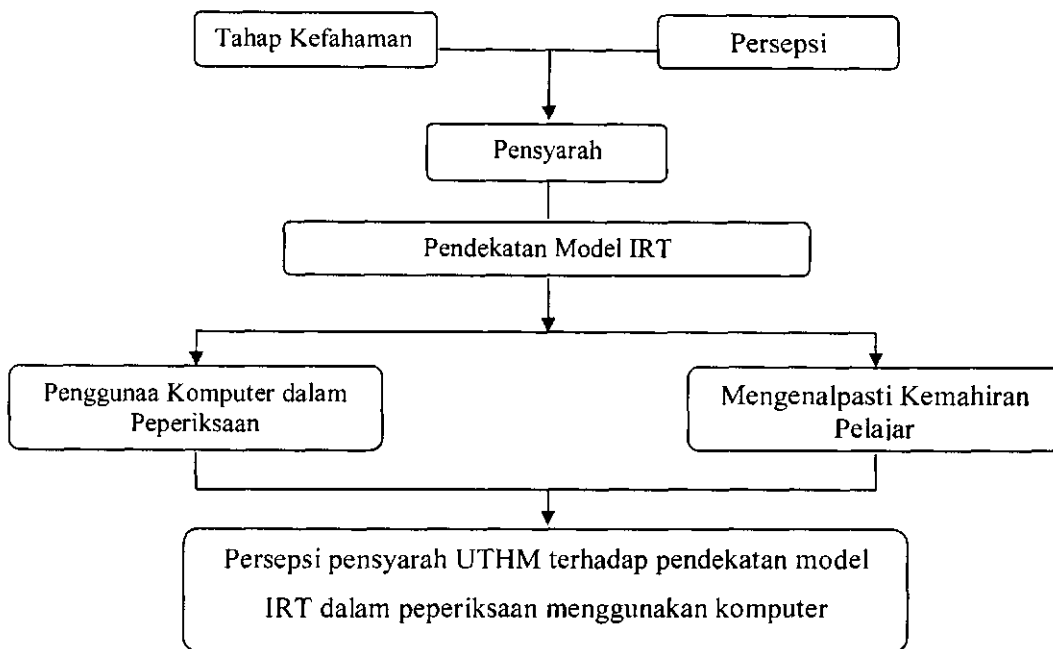
Anderson (2006) telah menjalankan kajian terhadap peperiksaan konvensional dan peperiksaan berdasarkan model IRT. Di mana pendekatan konvensional mempunyai sejarah penggunaannya yang lama dan kaedah penilaian dan pemarkahannya adalah mudah. Manakala IRT pula menggunakan pendekatan baru yang logik dan munasabah berasaskan model matematik dalam membuat penilaian dan pemarkahan terhadap pelajar. Iaitu kesukaran item (soalan) dalam peperiksaan dan tahap kebolehan pelajar mempunyai hubungan dan perkaitan dalam menghasilkan pelajar yang berkebolehan dan berkemahiran tinggi. Ini bermakna terdapat perbezaan yang ketara diantara sistem konvensional dengan IRT di dalam kaedah pengukuran dan penilaian pelajar.

Melalui pendekatan model IRT di dalam peperiksaan yang menggunakan komputer, ia merupakan satu alternatif untuk menggantikan peperiksaan konvensional kepada peperiksaan menggunakan teknologi komputer. Menurut Christina (2003) IRT merupakan satu model yang mempunyai keupayaan untuk menyelesaikan permasalahan di dalam peperiksaan dan pemarkahan. IRT menguruskan set soalan yang disimpan di dalam bank item dan diberi kepada pelajar mengikut kepada tahap kemampuan dan kebolehan seseorang pelajar itu menguasai sesuatu kemahiran dari set soalan yang mudah meningkat ke soalan yang lebih sukar untuk dijawab. Justeru itu, dalam kajian ini pengkaji melihat sejauhmanakah persepsi pensyarah-pensyarah di UTHM khususnya terhadap peperiksaan menggunakan komputer berdasarkan model IRT dari aspek kefahaman tentang konsep dan kebolehlaksanaannya di dalam sistem peperiksaan.

SKOP KAJIAN

Kajian ini ditumpukan kepada pendekatan model IRT dengan item pelbagai pilihan sahaja (*multiple choice item*). Kajian memberi tumpuan kepada persepsi pensyarah UTHM terhadap peperiksaan menggunakan komputer dengan pendekatan model IRT dalam mengenalpasti tahap kemahiran dan kebolehan pelajar dalam sesuatu bidang. Responden kajian terdiri daripada pensyarah di UTHM sahaja. Pemilihan responden adalah secara rawak di Fakulti Teknologi Maklumat & Multimedia, Fakulti Pengurusan Teknologi dan Fakulti Pendidikan Teknikal.

KERANGKA KONSEP



Rajah 1: Kerangka Konsep

Sumber : Adaptasi daripada James H.McMillan (2001) & Johnson And Johnson (2004)

Merujuk kepada Rajah 1, penilaian terhadap peperiksaan menggunakan komputer berdasarkan teori IRT akan dilakukan terhadap pensyarah di Fakulti Teknologi Maklumat & Multimedia, Fakulti Pengurusan Teknologi dan Fakulti Pendidikan Teknikal. Selain itu, aspek-aspek seperti tahap kefahaman dan persepsi pensyarah menjadi faktor penting terhadap kebolehlaksanaan peperiksaan menggunakan komputer berdasarkan teori IRT.

METODOLOGI KAJIAN

Kajian berbentuk kuantitatif ini berdasarkan kategori kuantiti seperti menggunakan angka, skor dan kekerapan. Dari segi teknik, data hasil daripada soal selidik berbentuk kuantitatif. Kajian menggunakan borang soal selidik dan rakaman contoh prototaip sistem peperiksaan menggunakan komputer sebagai instrumen kajian. Populasi kajian terdiri daripada pensyarah di Fakulti Pendidikan Teknikal, Fakulti Teknologi Maklumat & Multimedia dan Fakulti Pengurusan Teknologi, UTHM. Kaedah persampelan yang digunakan ialah secara rawak mudah. Sampel pensyarah adalah terdiri daripada 86 orang pensyarah daripada jumlah keseluruhan pensyarah di tiga fakulti yang berkenaan. Data dari 86 orang responden dikumpulkan dan dianalisis berdasarkan persoalan-persoalan kajian. Pengkaji menggunakan analisis deskriptif yang merujuk kepada analisis data yang diperolehi dari hasil soal selidik yang kemudiannya dijadikan satu set data yang senang difahami dan mudah diinterpretasikan. Data kajian telah dikodkan dan dianalisis menggunakan SPSS 11.0 (*Statistical Package For The Social Science*) untuk menentukan min dan sisihan piawai. Hasil analisis digolongkan ke dalam tiga tahap min iaitu berdasarkan kepada julat yang ditetapkan

DAPATAN KAJIAN

Maklumat Sistem Peperiksaan Menggunakan Komputer (Penggunaan Komputer Di dalam Peperiksaan)

Di dalam bahagian ini terdapat dua perkara yang diuji iaitu Penggunaan Komputer Di Dalam Peperiksaan dan Mengenalpasti Kemahiran Pelajar. Persepsi dan kefahaman pensyarah diukur

berdasarkan julat interpretasi skor min seperti yang dipersembahkan. Jadual 1 menunjukkan dapatan kajian yang telah dianalisis bagi set soal selidik yang dijawab oleh pensyarah.

Jadual 1 : Skor Min dan Sisihan Piawai Persepsi Pensyarah Terhadap Penggunaan Komputer Di dalam Peperiksaan

Bil	Item	Skor Min	Tahap
1.	Penggunaan komputer dalam peperiksaan adalah lebih efisien.	3.47	Sederhana
2.	Membenarkan pelajar menjalani peperiksaan dimana-mana sahaja.	3.94	Tinggi
3.	Pelajar lebih suka menggunakan komputer.	3.69	Tinggi
4.	Penggunaan komputer dalam peperiksaan boleh menjimatkan masa pelajar.	3.50	Sederhana
5.	Peperiksaan menggunakan komputer adalah lebih berstruktur dan lebih teratur dari segi pelaksanaannya.	3.65	Sederhana
6.	Penggunaan komputer dalam peperiksaan adalah memudahkan pelajar.	3.55	Sederhana
7.	Pengiraan markah adalah lebih tepat.	4.17	Tinggi
8.	Pengiraan markah adalah lebih adil.	4.01	Tinggi
9.	Komputer merupakan satu teknik membantu pelajar membuat keputusan bergantung kepada keupayaan pelajar tersebut.	3.64	Sederhana
10.	Komputer membantu dalam mengenalpasti tahap kefahaman pelajar.	3.55	Sederhana
SKOR MIN KESELURUHAN		3.7163	Tinggi

Berdasarkan Jadual 1, di dapati pensyarah bersetuju bahawa “pengiraan markah adalah lebih tepat dan adil” dengan menggunakan komputer iaitu setiap satu memperolehi skor min 4.17 dan 4.01. Selain itu, pensyarah juga turut bersetuju bahawa “pelajar boleh menjalani peperiksaan di mana-mana sahaja” dan “pelajar lebih gemar menggunakan komputer” dengan skor min masing-masing 3.94 dan 3.69. Merujuk kepada julat interpretasi skor min dalam Jadual 1, item-item ini berada pada tahap yang tinggi.

Selain itu didapati item-item lain seperti “penggunaan komputer yang lebih efisien”, “menjimatkan masa pelajar”, “pelaksanaan yang berstruktur”, “memudahkan pelajar”, “teknik membantu membuat keputusan” dan “membantu dalam mengenalpasti tahap kefahaman pelajar” berada pada tahap sederhana dengan skor min 3.47, 3.50, 3.65, 3.55, 3.64 dan 3.55. Tahap sederhana menunjukkan bahawa pensyarah tidak begitu bersetuju dengan item-item ini dalam peperiksaan yang menggunakan komputer. Walaubagaimanapun, skor min keseluruhan bagi persepsi pensyarah terhadap penggunaan komputer dalam peperiksaan adalah pada tahap yang tinggi iaitu 3.7163.

Analisis : Maklumat Sistem Peperiksaan Menggunakan Komputer (Menenalpasti Kemahiran Pelajar)

Item-item dalam bahagian ini dibina berdasarkan ciri-ciri model IRT (*item response theory*) yang digunakan di dalam sistem peperiksaan.

Jadual 2 : Skor Min dan Sisihan Piawai Kefahaman Pensyarah Terhadap Ciri-Ciri Model IRT Dalam Mengenalpasti Kemahiran Pelajar

Bil	Item	Skor Min	Tahap
1.	Pensyarah dapat mensasarkan tahap kebolehan pelajar dengan spesifik.	3.60	Sederhana
2.	Pensyarah dapat menganggarkan kebolehan pelajar dengan tepat.	3.38	Sederhana
3.	Tindakbalas diagnostik boleh disediakan dengan cepat untuk setiap pelajar pada mana-mana soalan yang dijawab salah.	3.91	Tinggi
4.	Item yang disediakan bersesuaian dengan kemahiran pelajar.	3.63	Sederhana
5.	Pelajar tidak akan menjawab soalan yang terlalu mudah sahaja.	3.49	Sederhana
6.	Pelajar tidak akan menjawab soalan yang terlalu sukar sahaja.	3.55	Sederhana
7.	Soalan yang dikemukakan adalah secara rawak.	3.83	Tinggi
8.	Kebolehan pelajar dapat dikenalpasti dengan lebih spesifik.	3.59	Sederhana
9.	Dapat mengelakkan penipuan di kalangan pelajar.	3.28	Sederhana
10.	Pensyarah boleh mengenal pasti pelajar yang mahir dan pelajar yang tidak mahir.	3.40	Sederhana
11.	Tahap kebolehan dan keupayaan pelajar dalam sesuatu bidang dikenal pasti dengan soalan yang mudah hingga soalan yang lebih sukar.	3.73	Tinggi
SKOR MIN KESELURUHAN		3.5803	Sederhana

Merujuk kepada Jadual 2, pensyarah bersetuju bahawa “tindakbalas diagnostik boleh disediakan dengan cepat untuk setiap pelajar pada mana-mana soalan yang dijawab salah” (skor min 3.91), “soalan yang dikemukakan adalah secara rawak” (skor min 3.83) dan “tahap kebolehan dan keupayaan pelajar dalam sesuatu bidang dikenal pasti dengan soalan yang mudah hingga soalan yang lebih sukar” (skor min 3.73) dalam peperiksaan yang menggunakan komputer. Item-item tersebut menunjukkan skor min pada tahap yang tinggi. Manakala bagi item-item “pensyarah dapat mensasarkan tahap kebolehan pelajar dengan spesifik” (skor min 3.60), “pensyarah dapat menganggarkan kebolehan pelajar dengan tepat” (skor min 3.38), “item yang disediakan bersesuaian dengan kemahiran pelajar” (skor min 3.63), “pelajar tidak akan menjawab soalan yang terlalu mudah sahaja” (skor min 3.49), “pelajar tidak akan menjawab soalan yang terlalu sukar sahaja” (skor min 3.55), dan “pensyarah boleh mengenal pasti pelajar yang mahir dan pelajar yang tidak mahir” (skor min 3.40) menunjukkan skor min pada tahap sederhana.

Namun begitu, bagi item “dapat mengelakkan penipuan di kalangan pelajar” menunjukkan skor min paling rendah iaitu 3.28. Ini menunjukkan kebanyakan pensyarah kurang bersetuju bahawa penggunaan komputer semasa peperiksaan dapat mengelakkan penipuan di kalangan pelajar. Secara keseluruhannya, tahap kefahaman pensyarah terhadap ciri-ciri model IRT dalam mengenalpasti kemahiran pelajar dalam peperiksaan menggunakan komputer berada pada tahap yang sederhana dengan skor min keseluruhan yang diperolehi ialah 3.5803.

Jadual 3 : Peratus Cadangan Atau Komen

Item	Keputusan	Bilangan	Peratus
Adakah wajar peperiksaan menggunakan komputer dengan pendekatan model IRT diaplikasikan dalam sistem peperiksaan negara kita.	Wajar	22	25.6
	Tidak Wajar	26	30.2
	Lain-lain	38	44.2
	Jumlah	86	100.0

Merujuk kepada Jadual 3, dapatan kajian menunjukkan 22 orang pensyarah bersetuju mengatakan bahawa peperiksaan menggunakan komputer ini wajar diaplikasikan di negara kita. Seramai 26 orang pensyarah pula mengatakan tidak wajar sistem peperiksaan ini diaplikasikan di negara kita. Manakala seramai 38 orang pensyarah tidak memberikan maklum balas terhadap item ini. Ini menunjukkan perlunya lebih pendedahan dan pengetahuan berkaitan dengan pendekatan model IRT di dalam sistem peperiksaan.

Rumusan Dapatan Kajian

Hasil kajian menunjukkan bahawa persepsi pensyarah terhadap penggunaan komputer di dalam komputer dengan pendekatan model IRT berada pada tahap sederhana dan masih memerlukan pendedahan yang lebih mengenainya. Ini adalah berdasarkan maklum balas yang diterima dari pensyarah di UTHM. Dapatan kajian juga menunjukkan bahawa kebanyakan pensyarah bersetuju dan mengakui bahawa dengan kehadiran teknologi mampu untuk meningkatkan taraf pendidikan terutamanya di dalam peperiksaan. Namun demikian, pendekatan model IRT perlu didedahkan dengan lebih meluas lagi agar perlaksanaannya benar-benar memberikan impak yang positif dan mampu untuk mengukur kemahiran seseorang dengan lebih spesifik.

PERBINCANGAN

Secara keseluruhannya, hasil dapatan kajian menunjukkan bahawa persepsi pensyarah berada di tahap yang sederhana. Ini menunjukkan penerimaan pensyarah terhadap penggunaan komputer di dalam peperiksaan dengan pendekatan model IRT masih memerlukan kajian yang lebih meluas. Dapatan kajian mendapati majoriti responden pensyarah di tiga fakulti menunjukkan persepsi yang tinggi terhadap penggunaan komputer dalam peperiksaan. Dimana responden bersetuju bahawa penggunaan komputer sedikit sebanyak membantu dalam pelaksanaan peperiksaan. Ini dibuktikan berdasarkan nilai skor min keseluruhan yang telah diperolehi dan dinyatakan dalam Jadual 4.

Jadual 4 : Dapatan Analisis Skor Min Keseluruhan Berdasarkan Persoalan Kajian Pertama

Persoalan Kajian Pertama	Responden	Skor min keseluruhan	Rumusan
Apakah persepsi pensyarah UTHM terhadap penggunaan komputer dalam peperiksaan?	Pensyarah	3.7163	Maklum balas responden mengenai penggunaan komputer di dalam peperiksaan adalah pada tahap yang "tinggi".

Berdasarkan hasil dapatan kajian yang diperolehi, pensyarah berpendapat bahawa penggunaan komputer dalam peperiksaan adalah lebih efisien kerana komputer merupakan teknologi yang semakin berkembang di negara kita. Norizan dan Raja Maznah (2004) menegaskan penggunaan komputer secara berkesan dan efisien dalam sesuatu kurikulum adalah dengan membenarkan pelajar belajar mengaplikasikan kemahiran komputer dengan cara yang bermakna. Penggunaannya yang boleh menjimatkan masa dan memudahkan pelajar menjadikan ia lebih efektif dan berkesan. Para responden juga turut bersetuju bahawa pelajar yang menggunakan komputer boleh menjalani

peperiksaan di mana-mana sahaja tempat. Suphat (2004) menyatakan bahawa kebanyakan pelajar gemar menggunakan komputer kerana teknologinya yang mesra pengguna dan memberikan keseronokan kepada pelajar.

Dengan penggunaan komputer di dalam peperiksaan, pengiraan markah telah diprogramkan dengan menggunakan formula-formula tertentu menjadikan pengiraannya menjadi lebih tepat dan adil. Penggunaan kertas-pensil di dalam peperiksaan konvensional menghadapi beberapa permasalahan dari segi masa yang tidak cekap, pembaziran pelbagai sumber dan melibatkan faktor kemanusiaan semasa proses pemarkahan (Maizatul, Seri Dewi dan T.Y. Yang, 2004). Kenyataan tersebut mengukuhkan bahawa komputer membantu di dalam proses pengiraan markah dengan lebih baik berbanding manusia. Ini dibuktikan berdasarkan skor min yang tinggi daripada responden terhadap pengiraan markah.

Penggunaan komputer merupakan satu teknik membantu pelajar membuat keputusan bergantung kepada keupayaan pelajar tersebut serta membantu dalam mengenalpasti tahap kefahaman pelajar. Kedua-dua item ini mempunyai skor min yang sederhana menunjukkan pensyarah kurang bersetuju bahawa komputer membantu dalam mengenalpasti kemahiran dan tahap kefahaman pelajar. Keadaan ini sedikit sebanyak menunjukkan percanggahan pendapat dengan Brown (1997) yang mengatakan bahawa *computer adaptive* merupakan kaedah yang termaju di dalam penilaian dimana komputer memilih dan mengemukakan item-item ujian kepada pelajar berpandukan kepada tahap kebolehan seseorang pelajar itu. Reckase (1989) juga di dalam kajiannya berpendapat yang sama dengan Brown iaitu *computer adaptive* merupakan salah satu sub kategori pengujian berasaskan komputer yang melibatkan pemilihan ujian secara dinamik terhadap soalan-soalan untuk menguji kemampuan sebenar pelajar semasa ujian dan akhirnya menjadi satu kaedah yang amat berkesan.

Mengenalpasti Kemahiran Pelajar

Dapatan kajian yang telah dilakukan, didapati majoriti responden pensyarah di tiga fakulti menunjukkan persepsi yang sederhana terhadap sistem peperiksaan dengan pendekatan Model IRT dalam mengenalpasti kemahiran pelajar. Pengkaji mendapati para responden kurang bersetuju bahawa penggunaan komputer boleh mengenalpasti kemahiran pelajar. Ini dibuktikan berdasarkan nilai skor min keseluruhan yang telah diperolehi adalah ditahap yang sederhana dan dinyatakan dalam Jadual 5.

Jadual 5: Dapatan Analisis Skor Min Keseluruhan Berdasarkan Persoalan Kajian Kedua

Persoalan Kajian Kedua	Responden	Skor min keseluruhan	Rumusan
Apakah persepsi pensyarah UTHM terhadap pendekatan model IRT dalam mengenalpasti kemahiran dan kebolehan pelajar melalui peperiksaan menggunakan komputer?	Pensyarah	3.5803	Maklum balas responden mengenai mengenalpasti kemahiran pelajar menggunakan sistem peperiksaan adalah pada tahap yang "sederhana".

Wainer (1990) menyatakan bahawa model IRT menggunakan jawapan-jawapan soalan semasa ujian dan juga soalan ujian untuk mengesan tahap kesukaran sesuatu soalan. Ini bermaksud model IRT berupaya untuk mengasingkan atau mengklasifikasikan soalan yang mudah hingga soalan yang sukar.

Berdasarkan hasil dapatan kajian yang diperolehi, pensyarah bersetuju bahawa soalan-soalan yang dikemukakan adalah secara rawak dan tahap kebolehan dan keupayaan pelajar dikenalpasti dengan

soalan yang mudah hingga soalan yang sukar. Secara spesifiknya, jika pemeriksa memberikan satu soalan yang agak senang untuk pelajar, soalan yang seterusnya adalah sangat sukar. Menurut Wainer (1990), pendekatan ini berpunca dari kesedaran mengenai kebolehan individu jika soalan itu terlalu mudah ataupun sukar. Pemeriksa akan mengetahui tentang kebolehan seseorang calon tersebut pada sesuatu tahap. Oleh itu, soalan pertama yang akan diberikan kepada pelajar biasanya adalah berbentuk sederhana sukar. Jika pelajar itu berjaya menjawab dengan betul, soalan yang seterusnya adalah lebih sukar. Walau bagaimanapun, jika pelajar tidak dapat menjawab dengan betul, satu soalan yang lebih mudah akan diberikan secara bergilir-gilir. Setelah itu, tindakbalas diagnostik boleh disediakan dengan cepat untuk setiap pelajar pada mana-mana soalan yang dijawab salah. Kenyataan ini disetujui oleh responden dengan nilai skor min yang agak tinggi berbanding item yang lain. Komputer akan mengesan dibahagian mana pelajar mempunyai masalah. Kaedah ini akan membantu pelajar dalam mengenalpasti dimana kelemahan pelajar dan apakah kaedah terbaik untuk pelajar berkenaan mengatasi masalah tersebut.

Pensyarah memberikan maklum balas yang menunjukkan mereka kurang setuju terhadap beberapa ciri-ciri yang terdapat dalam model IRT dalam mengenalpasti kemahiran pelajar. Ini dibuktikan melalui nilai skor min yang sederhana pada beberapa item. Antaranya adalah pensyarah dapat mensasarkan tahap kebolehan pelajar dengan spesifik dan tepat. Wainer (1990) menyatakan bahawa melalui pendekatan model IRT sesuatu kemahiran pelajar itu dapat dikenalpasti berdasarkan item-item yang disediakan kepada pelajar adalah menepati dan bersesuaian dengan kemahiran pelajar. Kenyataan ini sedikit sebanyak menunjukkan percanggahan pendapat dengan responden yang kurang bersetuju terhadap pemilihan item di dalam model IRT. Ini menunjukkan responden masih memerlukan maklumat dan pengetahuan yang luas mengenai ciri-ciri yang terdapat dalam model IRT.

Selain daripada itu, kebolehan IRT adalah mengesan penipuan dikalangan pelajar yang menjalani peperiksaan menggunakan komputer. Algoritma di dalam IRT dibangunkan bagi membolehkan pelajar yang menipu di dalam peperiksaan tidak diberi peluang sama sekali. Ini bertujuan agar IRT tidak melakukan kesilapan dalam mengklasifikasikan kemahiran dan kebolehan pelajar. Namun demikian, keadaan ini tidak disetujui oleh para responden. Ini dibuktikan melalui nilai skor min yang sederhana menunjukkan para responden kurang bersetuju bahawa peperiksaan berkomputer boleh mengelakkan penipuan dikalangan pelajar. Kebanyakan para responden meragui ciri-ciri dalam IRT ini disebabkan oleh sikap pelajar hari ini yang suka meniru semasa peperiksaan. Ini jelas menunjukkan keupayaan IRT dalam mengesan penipuan dikalangan pelajar masih belum dapat dibuktikan kepada para responden.

Persepsi Pensyarah Terhadap Pendekatan Model IRT

Dapatan kajian yang dijalankan, mendapati majoriti responden pensyarah di tiga fakulti menunjukkan persepsi yang sederhana terhadap pendekatan model IRT dalam peperiksaan menggunakan komputer. Para responden kurang bersetuju dari segi kebolehlaksanaan sistem peperiksaan ini diaplikasikan di dalam sistem peperiksaan negara. Ini dibuktikan berdasarkan nilai skor min keseluruhan yang telah diperolehi dan dinyatakan dalam Jadual 6.

Jadual 6: Dapatan Analisis Skor Min Keseluruhan Berdasarkan Persoalan Kajian Ketiga

Persoalan Kajian Ketiga	Responden	Skor min keseluruhan	Rumusan
Apakah persepsi pensyarah UTHM terhadap pendekatan model IRT dalam peperiksaan menggunakan komputer?	Pensyarah	3.5788	Maklum balas responden mengenai persepsi pensyarah secara keseluruhan terhadap pendekatan model IRT dalam peperiksaan menggunakan komputer berada ditahap yang "sederhana".

Dalam sistem pendidikan di luar negara, telah banyak kajian-kajian yang dilakukan yang membuktikan keberkesanan peperiksaan menggunakan komputer berdasarkan model IRT dalam menilai tahap kebolehan dan kemahiran pelajar. Antaranya adalah "*Computerized Adaptive Testing In Computer Science : Assessing Student Programming Abilities*" (Angel Syang dan Nell B.Dale, 1993), "*Classical Test Theory Or Item Response Theory : The Swedish Experience*" (Christina Stage, 2003), "*Considerations In Developing Or Using Second/foreign Language Proficiency Computer-adaptive Tests*" (Patricia A. Dunkel, 1999). Kajian-kajian ini menunjukkan bahawa peperiksaan menggunakan komputer berdasarkan model IRT dapat dilaksanakan dengan jayanya di luar Negara.

Namun demikian, perlaksanaannya di negara kita masih memerlukan penyelidikan yang lebih mendalam serta dari segi infrastruktur dan tahap kesediaan pensyarah dan pelajar itu sendiri sebelum diaplikasikan dalam sistem peperiksaan negara kita. Penerimaan bagi semua pihak amatlah penting bagi memastikan keberkesanannya dalam mengenalpasti kemahiran pelajar. Ini dibuktikan berdasarkan skor min yang berada ditahap sederhana menunjukkan pensyarah masih memerlukan maklumat yang lengkap terhadap pendekatan model IRT.

RUMUSAN

Majoriti responden mengakui mempunyai pengetahuan dan tahap kefahaman yang tinggi di dalam teknologi komputer. Namun demikian, tahap kefahaman responden terhadap ciri-ciri model IRT masih lagi ditahap yang rendah dan memerlukan maklumat yang lebih dalam sektor pendidikan. Dapatan hasil kajian menunjukkan tidak semua pensyarah memahami konsep dan ciri-ciri model IRT yang diterapkan di dalam sistem peperiksaan yang menggunakan komputer. Ini adalah berdasarkan kepada tahap kefahaman dan persepsi pensyarah yang rendah dalam sistem peperiksaan yang mengandungi ciri-ciri model IRT. Dari segi pelaksanaan atau kebolehlaksanaan sistem peperiksaan yang menggunakan komputer dengan pendekatan model IRT memerlukan lebih penelitian dan penyelidikan dari segi infrastruktur dan kemampuan model IRT dalam mengenalpasti kemahiran pelajar. Ini disebabkan oleh kemudahan sokongan yang lemah dan ketidakpercayaan pensyarah terhadap sikap pelajar itu sendiri.

Fenomena penipuan dalam kalangan pelajar juga merupakan suatu senario yang sangat diragui dalam melaksanakan sistem peperiksaan menggunakan komputer ini. Dengan pelaksanaan yang dilihat seperti kelemahan pemantauan serta sikap pelajar itu sendiri menyebabkan persepsi pensyarah terhadap perlaksanaannya berada ditahap yang rendah. Ciri-ciri yang terdapat dalam model IRT sebenarnya telah dilengkapkan dengan model yang boleh mengesan pelajar yang meniru di dalam peperiksaan. Namun demikian, tanpa bukti dan kajian mengenainya ia masih belum boleh dilaksanakan lagi.

Bagi meningkatkan tahap pengetahuan dan kefahaman terhadap pendekatan model IRT Pensyarah perlu didedahkan dengan lebih terperinci tentang ciri-ciri pendekatan model IRT di dalam peperiksaan agar dapat memahami konsep dan pelaksanaannya dalam mengenalpasti kemahiran dan kebolehan pelajar di dalam sesuatu bidang.

RUJUKAN

- Angel Syang, Nell B.Dale (1993). *Computerized Adaptive Testing In Computer Science : Assessing Student Programming Abilities*. The University of Texas.
- Brown, J.D. (1997). *Computers In Language Testing : Present Research And Some Future Directions*. Journal Of Language Learning & Technology.
- Christina Stage (2003). *Classical Test Theory Or Item Response Theory : The Swedish Experience*. Chile. Centro de Estudios Publicos.
- Eva Jettmar & Clifford Nass (2002). *Adaptive Testing: Effects on User Performance*. Stanford University. Proceeding Paper.
- James H.McMillan. (2001). *Classroom Assessment Principles And Practice For Effective Instruction*. Second Edition. United States Of America. A Pearson Education Company.9 – 18.
- John O. Anderson (2006). *Does Complex Analysis (IRT) Pay Any Dividends In Achievement Testing?*. University Of Victoria. Present Paper.
- Johnson, D.W. & Johnson,R.T. (1994). *Learning Together And Alone : Cooperative, Competitive and Individualistic Learning*. Boston : Allyn and Bacon.
- Marie Wiberg.(2004). *Classical Test Theory vs. Item Response Theory*.Umea University.
- Marie H. Wiberg (2003). *Computerized Achievement Tests - sequential and fixed length tests*. Umea University.
- Patricia A. Dunkel. (1999). *Considerations In Developing Or Using Second/Foreign Language Proficiency Computer-Adaptive Tests*. Journal Of Language Learning & Technology.
- Suphat (2004). *Computerized Test/Item Banking And Computerized Adaptive Testing For Teachers And Lecturers*. Language Institute Chulalongkorn University.
- Tom Kubiszyn, Gary Borich (2003). *Educational Testing And Measurement Classroom Application And Practice*. Seventh Edition. United States Of America. John Wiley & Sons,INC. 57 – 74.
- Wainer, H.(1990). *Computer Adaptive Testing : A Primer*. Hillsdale, NJ:Lawrence Erlbaum Associates.