



## Keupayaan Teori dan Pelaksanaan Pentaksiran dalam Pembelajaran

Arsaythamby Veloo\*

*Pusat Pengajian Pendidikan dan Bahasa Moden, Universiti Utara Malaysia*

\*Corresponding author; email: arsay@uum.edu.my

### ABSTRAK

Kertas konsep ini adalah bertujuan untuk mengenal pasti aspek-aspek ujian dari segi teori ujian, penilaian dalam pengajaran dan pembelajaran, hubungan objektif pengajaran dan penilaian dan jenis-jenis ujian. Teori ujian menemukan bagaimana teori ini dikaitkan dengan konsep-konsep ujian seperti maksud ujian, pengukuran dan penilaian itu sendiri. Teori ujian juga dikaitkan dengan bidang psikologi dan pendidikan. Di samping itu penilaian juga digunakan untuk mengetahui keberkesanan pengajaran guru dan penguasaan pelajar di sekolah. Dalam pengajaran dan pembelajaran yang berkesan maka hubungan antara objektif, pengajaran dan pembelajaran dan penilaian adalah penting dan ketiga-tiga aspek ini mempunyai hubungan yang saling berkaitan antara satu sama lain. Kertas ini juga membincangkan jenis-jenis ujian dalam pendidikan berdasarkan ujian objektif dan subjektif yang kerap kali dikendalikan di sekolah.

*Katakunci: pengajaran, pengukuran, penilaian, pembelajaran, ujian*

## Theoretical Ability and Implementation of Assessment in Learning

### ABSTRACT

This conceptual paper is aimed at identifying testing aspects based on testing theory, teaching and learning evaluation, objectives of teaching-evaluation relationship and types of testing. The testing theory find how the theory is associated with the testing concepts such as the meaning of testing, measurement and evaluation. Testing theory is also related to the fields of psychology and education. In addition, evaluation is also used to gauge the effectiveness of teacher's teaching and student's mastery in school. In effective teaching and learning, the relationship between the objectives, teaching and learning and evaluation are important and these three aspects have close relationship between one and the other. This paper discusses the types of testing in education based on the objective and subjective tests which are often conducted in schools.

*Keywords: teaching, measurement, evaluation, learning, testing*

### Pengenalan Teori Ujian

Bagi memahami keperluan teori ujian ini adalah penting untuk memahami tentang sifat asas pengukuran, gagasan, dan ujian psikologikal. Weitzenhoffer (1951) menjelaskan pengukuran sebagai "satu operasi yang dilaksanakan ke atas dunia fizikal oleh seseorang pemerhati". Bagi Stevens (1946) pengukuran ialah "pelbagai tugas kepada objek atau peristiwa berdasarkan peraturan." Lord dan Novick (1968) dan Torgerson (1958) menambahkan ketepatan definisi Steven dengan memberi perhatian bahawa tiada

pengukuran diaplikasikan kepada sifat objek melainkan daripada objek itu sendiri. Apabila seseorang ahli kimia fizikal mengukur jisim molekul satu sebatian atau seorang ahli biologi menentukan kandungan bakteria dalam setitik air kolam, ini adalah pengukuran sifat khusus ke atas objek itu dalam penyiasatan. Seorang ahli psikologi sekolah tidak mengukur seorang kanak-kanak tetapi sifat-sifat fizikal kanak-kanak tersebut seperti ketinggian atau berat atau satu sifat psikologi seperti perkembangan perbendaharaan kata, kematangan sosial, atau pengetahuan tentang asas penambahan (Crocker & Algina, 1986).

Psikologi didefinisikan secara tradisional sebagai satu bidang kajian tingkah laku dan ciri yang mengkategorikan tingkah laku seseorang individu di rumah, tempat kerja, sekolah atau latar belakang sosial kebiasaannya dikenali sebagai ciri psikologi (sifat psikologi). Tidak seperti ciri fizikal, ciri psikologi seseorang individu tidak boleh diukur secara langsung seperti ketinggian atau berat. Ciri psikologi adalah konstruk. Ia adalah konsep andaian – produk daripada imaginasi saintifik oleh ahli sains sosial yang mana ia berusaha membangunkan teori untuk menjelaskan tingkah laku manusia. Kewujudan gagasan sebegini tidak boleh disahkan sepenuhnya. Oleh itu, tahap di mana gagasan psikologi mengkategorikan seseorang individu hanya boleh dirumuskan daripada pemerhatian ke atas tingkah laku seseorang (Crocker & Algina, 1986).

Proses pembinaan gagasan dan bagaimana ia memberi panduan ke arah pengukuran sifat-sifat dapat diterangkan melalui contoh berikut. Katakan seseorang ahli pembangunan psikologi, berkhidmat di tadika, mendapati pelbagai aktiviti permainan yang mana sesetengah kanak-kanak cuba berulang kali untuk mengawal atau mengarah rakan-rakan semasa bermain. Setelah memerhati tingkah laku ini secara konsisten untuk satu jangka waktu dan dalam konteks yang berbeza untuk individu-individu yang sama, ahli psikologi boleh mula melabelkan tingkah laku ini sebagai “dominasi sosial”. Ahli psikologi ini telah mencipta satu teori gagasan yang merangkumi beberapa siri kelakuan yang sama (Crocker & Algina, 1986). Namun, mencipta konstruk tidak sama seperti mengukur konstruk. Sebelum pengukuran konstruk dibuat adalah penting untuk megukuhkan beberapa peraturan kesamaan antara konstruk teori dan tingkah laku yang boleh diperhati yang merupakan penentu yang sah bagi konstruk itu. Proses ini dikenali sebagai *pengukuhan tafsiran*. Dalam contoh di atas untuk mengukur kombinasi sosial, ahli psikologi perlu mengkhususkan jenis tingkah laku dalam ruang permainan di tadika yang dianggap “mendominasi.” Kemudian ahli psikologi perlu membina satu rancangan untuk memperolehi contoh tingkah laku sebegini dalam satu situasi standard dan merekod pemerhatian untuk setiap kanak-kanak dalam format yang standard. Ini memerlukan pembinaan satu instrumen atau ujian, untuk membentuk dominasi sosial.

Ujian boleh didefinisikan sebagai prosedur standard untuk memperolehi satu sampel ingkah laku daripada satu domain yang spesifik. Seperti yang diguna dalam teks ini, konsep ujian merujuk kepada prosedur untuk memperolehi satu sampel pencapaian maksimum seseorang individu (seperti yang dilambangkan oleh bakat atau ujian pencapaian di mana orang yang diuji di arahkan untuk membuat yang terbaik) atau satu sampel pencapaian tipikal seorang individu (seperti soalan kaji selidik atau inventori di mana responden melaporkan perasaan tipikal mereka, kelakuan, minat atau reaksi kepada situasi). Pendekatan lain dalam persampelan pencapaian tipikal ialah satu jadual standard dan senarai kelakuan yang boleh diguna oleh seseorang pemerhati yang merekod kelakuan yang ditonjolkan oleh subjek dalam keadaan semulajadi (Crocker & Algina, 1986). Apabila seseorang pakar psikologi menyediakan satu senarai tingkah laku untuk disemak atau dinilai oleh seorang pemerhati dengan menggunakan jadual yang ditentukan terlebih dahulu, apabila seorang guru kimia menulis satu set item aneka pilihan menggunakan elemen dari pelbagai peringkat aktiviti dalam jangkamasa tertentu. atau apabila seseorang kaunselor menilai kerisauan pelanggan dalam satu senarai situasi sosial. Setiap mereka dalam situasi masing-masing telah memikirkan satu ujian di dalam konteks yang luas. Kita akan menggunakan konsep ujian dalam konteks ini.

Pengukuran ciri psikologi berlaku apabila satu nilai kuantitatif ditugaskan atau ditentukan kepada sampel tingkah laku yang dikumpul melalui ujian. Dalam erti kata lain satu pengukuran telah diambil apabila ahli

psikologi selaraskan bilangan tingkah laku “dominasi sosial” ke atas senarai semak yang ditunjukkan oleh seorang kanak-kanak dalam tempoh masa lima minit ketika pemerhatian atau semasa guru kimia mengira bilangan item yang dijawab oleh seorang pelajar dengan betul dan merekodkan jumlah skor (Crocker & Algina, 1986). Daripada pengukuran tentang tingkah laku yang diperhatikan penguji membuat inferens jumlah tingkah laku individu berdasarkan teori gagasan. Kenapa ahli psikologi dan para penyelidik cuba untuk mengukur gagasan yang tidak dapat diukur secara langsung? Dalam pendidikan dan psikologi, sesuatu gagasan memberikan satu kaedah yang efisien dan sesuai untuk melabel beberapa tingkah laku yang sama (Crocker & Algina, 1986; Thorndike, 2001).

Kebanyakan percubaan untuk memerhati fenomena kelakuan kompleks akan menjadi huru hara dan mengelirukan. Bayangkan sejumlah perkataan, perlakuan dan interaksi yang boleh berlaku dalam satu hari di sebuah kelas. Cronbach (1970), mungkin menyalahafsirkan keadaan ini apabila ia menyatakan bahawa pengkajian aktiviti dalam bilik darjah adalah sesukar seperti mengkaji fenomena ribut taufan. Seseorang pemerhati yang cuba merekod dan menggambarkan tingkah laku setiap individu dalam situasi ini akan ditenggelami dengan tugas yang mesti dilakukan (Thorndike, 2001). Tambahan pula sebarang percubaan untuk meramal, mengawal atau menentukan punca setiap tingkah laku adalah sia-sia. Namun dengan penggunaan konstruk, pemerhati boleh mula mengklasifikasikan dan mengumpulkan contoh tingkah laku yang sama dan berinteraksi dalam tempoh masa yang telah disaksikan. Konstruk seperti ini ialah asas kepada teori tingkah laku manusia.

Pada tahap yang mudah satu teori psikologi adalah satu pernyataan tentang kemungkinan hubungan antara dua konstruk psikologi ataupun di antara satu konstruk dan satu fenomena yang boleh diperhatikan dalam satu siri urutan. Akhirnya melalui penyiasatan empirikal dan pengesahan teori seperti ini adalah mungkin untuk meramal dan mengawal corak atau tren sesetengah tingkah laku. Bagi mencapai tujuan ini adalah perlu untuk mengkuantifikasikan pemerhatian terhadap tingkah laku yang diwakili oleh konstruk yang dikaitkan dengan teori (Crocker & Algina, 1986).

## **PENILAIAN DALAM PENGAJARAN DAN PEMBELAJARAN**

Penilaian merupakan satu proses yang dilakukan untuk mengetahui keberkesanan pengajaran guru dan penguasaan pelajar dalam sesuatu kursus peringkat sekolah. Penilaian memainkan peranan yang penting dalam penambahbaikan pengajaran dan pembelajaran guru dalam bilik darjah.

Penilaian dalam pendidikan adalah sangat penting kepada semua pihak iaitu pelajar, ibu bapa dan para guru. Oleh itu penilaian perlu dilakukan dengan sempurna dan berkesan. Guru perlu merangka objektif dan kandungan mata pelajaran atau kursus yang hendak diajar sebelum strategi dan kaedah pengajaran dilaksanakan. Dalam proses pengajaran dan pembelajaran guru, penilaian terhadap pelajar adalah berdasarkan penilaian formatif dan sumatif. Dalam penilaian formatif, pengajar selalunya menggunakan beberapa kaedah atau ujian untuk mengukur kebolehan pelajar dalam menguasai kursus yang sedang diikuti (Thorndike, 2001). Antara penilaian formatif yang selalunya dikendalikan adalah kuiz, tugas, pembentangan dan sebagainya. Manakala ujian sumatif hanya melibatkan peperiksaan akhir pada setiap kursus. Oleh itu seseorang guru perlu menentukan kaedah yang digunakan untuk menentukan kebolehan pelajar dalam sesuatu kursus yang sedang diikuti. Seseorang guru yang mengajar pelajar harus mengetahui konsep-konsep pengujian, pengukuran dan penilaian dalam pendidikan sebelum memahami penilaian secara holistik (Kubiszyn & Borich, 2003).

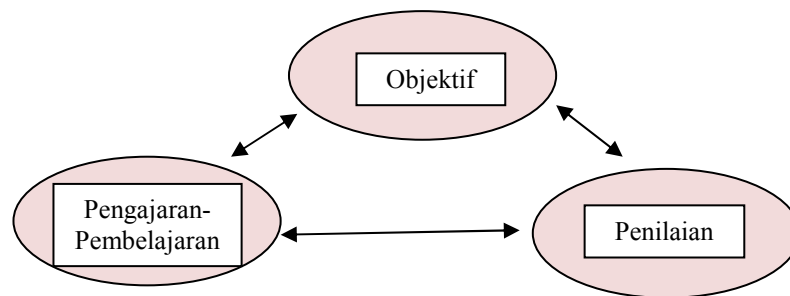
Pengukuran penguasaan pelajar menjadi sempurna sekiranya guru menggunakan pelbagai jenis ujian untuk memperolehi skor sesuatu kursus dalam bentuk 100%. Bagi mengukur penguasaan pelajar, guru perlu menggunakan alat ukur mengikut kandungan kursus. Alat ukur yang dimaksudkan disini boleh jadi kuiz, tugas, pembentangan atau peperiksaan yang terdiri daripada penilaian formatif dan sumatif.

Sesuatu jenis ujian yang digunakan untuk mengukur mestilah memenuhi beberapa kriteria atau ciri-ciri ujian iaitu kesahan, kebolehpercayaan, keobjektifan dan juga kebiasaan. Di samping itu penilaian merupakan komponen ujian dan pengukuran yang menghasikan skor untuk menterjemahkan penguasaan pelajar dalam bentuk nilai numerik mengikut hasil pembelajaran yang tertera dalam kursus yang dikendalikan (Popham, 2000).

## HUBUNGAN OBJEKTIF PENGAJARAN DAN PENILAIAN

Objektif pengajaran memainkan peranan penting dalam pengajaran guru di dalam bilik darjah. Dalam setiap model pengajaran terdapat tiga elemen yang penting iaitu objektif, pengajaran dan Pembelajaran, dan penilaian. Objektif pengajaran memberikan panduan dan matlamat kepada para guru dan pelajar mengenai pengajaran dan pembelajaran. Berdasarkan objektif pengajaran, para pengajar dapat mengendalikan kaedah pengajaran dan bahan pengajaran yang sesuai untuk mencapai objektif yang ditentukan (Pang & Denis Lajium, 2010). Secara ringkasnya objektif pengajaran penting dalam pengajaran dan penilaian. Hubungan ketiga-tiga elemen digambarkan dalam Rajah 1.

**Rajah 1**  
**Hubungan antara Objektif, Pengajaran-pembelajaran dan Penilaian**



## JENIS-JENIS UJIAN DALAM PENDIDIKAN

Pada prinsipnya ujian perlu dibina dengan sistematik dan jelas agar matlamat untuk memperoleh bukti markat yang sah dan dipercayai bagi tujuan penilaian pengajaran dicapai. Guru perlu mengaplikasikan prinsip dan teknik penggubalan dan penyediaan item supaya ujian objektif dan subjektif yang dihasilkan itu adalah bermutu (Crocker & Algina, 1986). Bagi tujuan penggubalan item ujian, kefahaman terhadap pengertian kata tugas yang akan digunakan bagi penggubalan item bagi ujian objektif mahupun subjektif perlu dipastikan supaya item-item yang digunakan bukan sahaja jelas malah tepat mengukur sesuatu objektif pengajaran. Dengan memahami persamaan dan perbezaan serta kekuatan dan kelemahan kedua-dua teknik ujian serta kata tugas yang sesuai digunakan akan meningkatkan ketekalan ujian. Oleh itu, guru seharusnya berupaya menggubal sesuatu item ujian berasaskan kesesuaiannya dengan penilaian yang akan digunakan untuk menilai keberkesanan pengajaran dalam bilik darjah (Kubiszyn & Borich, 2003).

Ujian yang dikendalikan mengukur jenis gerak balas yang diinginkan dari pelajar. Ujian yang hanya memerlukan pelajar memilih jawapan yang telah diberi dikenali sebagai ujian objektif. Jenis jawapan dipilih memerlukan pelajar memilih jawapan daripada pilihan jawapan yang telah diberikan. Guru bukan sahaja perlu menyediakan item malah perlu menyediakan jawapan yang paling tepat bagi item berkenaan beserta dengan pengganggu. Di antara ujian jenis jawapan dipilih yang sering digunakan ialah jenis soalan isi tempat kosong, salah benar, pemadanan, aneka pilihan dan aneka pilihan gabungan. Guru perlu

memilih jenis item yang bersesuaian dengan keperluan objektif pengajaran dan bertujuan untuk mengukur kebolehan pelajar mengikut sukatan pelajaran (Popham, 2000).

### **Ujian Objektif**

Pada umumnya ujian objektif mempunyai objektiviti yang tinggi kerana jawapan bagi setiap pilihan item diberikan hanya satu sahaja jawapan yang betul. Oleh itu untuk memeriksa jawapan ini tidak perlu pengetahuan tentang subjek berkenaan. Sesiapa sahaja boleh memeriksa kertas jawapan ini termasuk mesin. Walaubagaimana pun ujian ini juga dikatakan tidak benar-benar objektif kerana terdapat unsur-unsur subjektif dalam memilih jenis item ujian yang tidak sesuai. Pemarkatan ujian ini juga sangat objektif sehingga jenis pelbagai pilihan sering dikenali sebagai soalan objektif. Analisis item dilakukan berdasarkan pekali diskriminasi dan kesukaran (Embretson, & Hershberger, 1999). Item ujian objektif terdiri daripada dua bahagian iaitu stem dan opsyen. Stem boleh dibina berdasarkan soalan atau pernyataan. Manakala bahagian opsyen terdiri daripada jawapan dan distraktor (Kubiszyn & Borich, 2003). Bagi ujian objektif hanya terdapat satu sahaja jawapan yang betul dan yang lain adalah jawapan yang salah. Kadang-kadang terdapat banyak jawapan tetapi dipilih jawapan yang terbaik. Ujian objektif terdiri daripada lima jenis ujian iaitu isi tempat kosong, salah benar, padanan, aneka pilihan dan aneka pilihan gabungan. Setiap jenis ujian ini diterangkan secara mendalam seperti berikut:

#### **i. Item Isi Tempat Kosong**

Jenis item ini memerlukan ketelitian yang mendalam untuk menjawab dan jawapannya perlu diketahui dengan tepat. Garisan atau ruang kosong yang disediakan untuk jawapan isi tempat kosong adalah sama besar atau panjang. Pelajar perlu mempunyai kemahiran dalam memilih jawapan yang tepat berdasarkan kenyataan yang diberi. Adakalanya jenis item ini memerlukan dua jawapan dengan mengisi dua tempat kosong.

Item isi tempat kosong menghendaki pelajar melengkapkan ayat atau rumusan sesuatu kenyataan. Isi tempat kosong ada dua kategori iaitu jenis jawapan telah diberi dan jawapan yang tidak diberi. Item isi tempat kosong mudah dibina dan ditulis. Item jenis ini juga boleh menampung lebih banyak kandungan berbanding dengan ujian aneka pilihan kerana jawapan item isi tempat kosong mudah ditulis. Di samping itu, soalan jenis ini juga berupaya mengelak tekaan kerana item jenis ini memerlukan daya ingatan yang tinggi (Embretson, & Hershberger, 1999).

Item jenis ini sesuai untuk mengukur ingatan, fakta, nama, tempat dan peristiwa khusus dan tidak sesuai untuk tingkah laku yang sukar diramal seperti pemikiran kreatif dan kritis. Di samping itu, item isi tempat kosong sangat sukar dimarkat sekiranya item tidak disertakan dengan pilihan jawapan yang betul. Kelemahan utama item isi tempat kosong ialah menggalak jawapan aras rendah yang tidak sesuai untuk objektif aras tinggi.

#### **ii. Item Betul-Salah**

Item jenis betul-salah menghendaki pelajar menyatakan sama ada sesuatu kenyataan yang dikemukakan betul atau salah. Item jenis ini mempunyai banyak kelebihan. Item jenis ini menguji banyak kandungan dan bahan pengajaran dan pembelajaran. Lagipun item jenis ini memerlukan masa yang sedikit untuk membina dan memerlukan masa pemarkatan jawapan yang pendek. Namun item jenis ini lebih menekankan kepada daya ingatan pelajar daripada mengembangkan daya intelek kerana terhad kepada jawapan betul atau salah dan mempunyai kebarangkalian tekaan yang tinggi. Item jenis ini mempunyai kebarangkalian penekanan yang tinggi iaitu sebanyak 50 peratus untuk menjawab dengan betul.

### **iii. Item Pemadanan**

Item jenis pemadanan menghendaki pelajar memadankan sesuatu ciri dengan gagasan, konsep atau prinsip yang berkaitan. Item bentuk pemadanan mudah dibina dan dimarkat serta sesuai untuk menguji kefahaman pelajar dalam mengaitkan antara fakta ciri yang dinyatakan dengan gagasan, konsep atau prinsip yang berkaitan. Item pemadanan juga mempunyai kebarangkalian tekaan jawapan benar yang rendah bergantung kepada bilangan pilihan yang diberi. Bagaimanapun, item pemadanan cenderung untuk bertanya maklumat yang remeh-temeh. Kesukaran juga sering berlaku apabila terdapat bilangan pilihan padanan yang banyak. Kebanyakan borang jawapan komersial mempunyai ruang jawapan yang terhad bagi sesuatu item.

Arahan bagi item pemadanan mestilah jelas. Item-item pada setiap set hendaklah serupa dari segi isi yang diuji. Susunan item perlu diteliti agar kenyataan ciri-ciri sesuatu pilihan yang biasanya menggunakan ayat yang lebih panjang disenaraikan dahulu sebelum mengemukakan pilihan jawapan yang biasanya menggunakan perkataan pendek. Para guru juga perlu mengelakkan bilangan kenyataan sama dengan bilangan pilihan jawapan bagi mengurangkan tekaan (Popham, 2000). Kebarangkalian tekaan boleh dikurangkan dengan memaklumkan kepada pelajar bahawa setiap pilihan yang diberi boleh digunakan lebih daripada sekali.

### **iv. Item Aneka Pilihan**

Item aneka pilihan menghendaki pelajar memilih jawapan yang paling tepat daripada beberapa pilihan jawapan bagi soalan yang dikemukakan. Justeru itu, item jenis pelbagai pilihan bukan sahaja mengandungi satu pilihan jawapan yang benar tetapi juga mengandungi pengganggu. Pengganggu adalah pilihan jawapan yang salah. Pengganggu dibina untuk mengelirukan pelajar. Walaupun salah, pengganggu perlu dibina dengan teliti supaya ianya memainkan peranan sebagai pengganggu yang berkesan dan seterusnya mengelakkan pemilihan jawapan yang dibuat secara tekaan.

Item aneka pilihan mempunyai banyak kelebihan. Item aneka pilihan boleh digunakan untuk menguji pelbagai tahap objektif pengajaran dari aras pengetahuan ke penilaian untuk domain kognitif hingga ke pelbagai aras domain afektif dan psikomotor (Embretson, & Hershberger, 1999). Ujian jenis ini membolehkan pelajar membezakan tahap ketepatan jawapan. Ujian aneka pilihan juga boleh menguji banyak kandungan pelajaran. Pemarkatan ujian ini juga sangat objektif sehingga jenis aneka pilihan sering dikenali sebagai ujian objektif. Analisis item seperti analisis diskriminasi dan analisis kesukaran mudah dilakukan.

Namun, ujian aneka pilihan juga mempunyai kelemahan. Pembinaan item aneka pilihan memerlukan masa dan tenaga yang banyak. Sekiranya item ujian pelbagai pilihan tidak ditulis dengan baik dan sempurna, kemungkinan terdapat beberapa pilihan jawapan yang betul. Kebarangkalian tekaan tinggi (25 peratus) sekiranya bilangan pilihan jawapan 4 dan menurun kepada 20 peratus sekiranya bilangan pilihan jawapan ditingkatkan kepada 5.

### **v. Item Aneka Pilihan Gabungan**

Item aneka pilihan gabungan merupakan item yang mempunyai gabungan beberapa jawapan yang betul sebagai jawapan. Manakala item aneka pilihan hanya mempunyai satu pilihan sebagai jawapan. Oleh itu item jenis aneka pilihan gabungan memerlukan pengetahuan yang tinggi terhadap sesuatu item untuk dijawab. Kerap kali jenis item ini dianggap sebagai sangat sukar dijawab oleh pelajar.

## Ujian Objektif

Ujian subjektif terdiri daripada tiga jenis item iaitu item jawapan pendek, item berstruktur dan item esei. Item jawapan pendek dikategorikan sebagai jenis ujian subjektif kerana pelajar perlu membina jawapan yang tepat berdasarkan kepada kefahaman sendiri. Item jenis ini adalah paling senang bagi pelajar kerana ia memerlukan jawapan aras pengetahuan dan kefahaman sahaja. Jenis Item ini mempunyai keobjektifan yang tinggi berbanding dengan item berstruktur dan esei (Popham, 2000). Item ini diberikan markah antara 2 hingga 4 dan senang diperiksa kerana jawapannya berdasarkan kepada fakta semata-mata. Ketiga-tiga jenis ujian dibincangkan seperti berikut

### i. Item Jawapan Pendek

Item jawapan pendek dikategorikan sebagai jenis ujian subjektif kerana pelajar perlu membina jawapan yang tepat berdasarkan kepada kefahaman sendiri. Item jenis ini adalah paling senang bagi pelajar kerana item jenis ini memerlukan jawapan aras pengetahuan dan kefahaman sahaja. Jenis item ini mempunyai keobjektifan yang tinggi berbanding dengan item berstruktur dan esei. item ini diberikan markah antara 2 hingga 4 dan senang diperiksa kerana jawapannya berdasarkan kepada fakta semata-mata.

### ii. Soalan Berstruktur

Item berstruktur ini adalah sangat popular berbanding dengan jenis item esei dan jawapan pendek terutamanya peringkat universiti. Item berstruktur senang diperiksa kerana memerlukan jawapan yang padat. Item jenis ini selalunya terdiri dari aras pengetahuan dan kefahaman. Skema markah pula senang disediakan bagi bentuk item ini dan juga jenis ini mudah diperiksa.

### iii. Item Esei

Item esei merupakan item yang paling gemar digunakan oleh pensyarah peringkat universiti. Ini kerana item jenis ini adalah sesuai untuk menguji aras yang lebih tinggi pada peringkat aras sintesis dan penilaian. Item esei mudah disediakan tetapi agak sukar untuk menyediakan skema pemarkahannya. Di samping itu, pemeriksa juga agak sukar untuk memeriksa jawapan ini dengan konsisten. Justeru, jenis item ini mempunyai keobjektifan yang rendah berbanding dengan jenis jawapan pendek dan berstruktur.

## KESIMPULAN

Penilaian yang dibuat terhadap seseorang pelajar adalah penting untuk menentukan kualiti pelajar yang dilahirkan hasil pengajaran dan pembelajaran di semua peringkat pengajian. Hasil penilaian juga sangat bergantung kepada jenis-jenis ujian yang dikendalikan di samping aspek kesahan dan kebolehpercayaan ujian.

## RUJUKAN

- Crocker, L., & Algina, J. (1986). *Introduction to classical & modern test theory*. New York: Holt, Rinehart & Winston.
- Cronbach, L. J. (1970). Review of on the theory of achievement test items. *Psychometrika*, 35, 509-511.
- Embretson, S. E., & Hershberger (1999). *The new rules of measurement: What every psychologist and educator should know*. London: Lawrence Erlbaum Associates.
- Kubiszyn, T. & Borich, G. D. (2003). *Educational testing and measurement: Classroom application and practice* (7<sup>th</sup> ed.). New York: John Wiley & Sons.

- Lord, F. M. & Novick, M. R. (1968). *Statistical theories of mental test scores*. Reading, Mass.: Addison-Wesley.
- Pang, V., & Denis Lajium (2010). *Penilaian dalam pendidikan* (2nd ed.). Universiti Malaysia Sabah, MAPIM.
- Popham, W. J. (2000). *Modern educational measurement: Practical guidelines for educational leaders* (3<sup>rd</sup> ed.). Boston: Allyn and Bacon.
- Stevens, S. S. (1946). On the theory of scales of measurement. *Science*, 103, 677-680.
- Thorndike, R. L. (2001). *Measurement and evaluation in psychology and education* (6<sup>th</sup> ed.). Merrill, Upper Saddle River, New Jersey: Columbus.
- Torgerson, W. S. (1958). *Theory and methods of scaling*. New York: John Wiley.
- Weitzenhoffer, A. M. (1951). Mathematical structures and psychological measurement. *Psychometrika*, 16, 387 – 406.