BAB II
TINJAUAN TEORITIK MENGENAI TES INTELIGENSI DALAM BIMBINGAN DAN KONSELING

A. Bimbingan dan Konseling dalam Mengembangkan Potensi Individu

1. Hakikat Layanan BK dalam Pengembangan Potensi

Bimbingan merupakan upaya fundamental pada setiap ikhtiar pendidikan, baik pendidikan formal, non-formal maupun informal. Dalam ketiga bentuk pendidikan tersebut, proses bimbingan (*guidance*) dipastikan selalu melekat di dalamnya. Berbeda dengan pengajaran, yang tidak selalu harus ada di dalam setiap bentuk pendidikan tersebut.

Bimbingan pada hakikatnya merupakan proses memfasilitasi pengembangan nilai-nilai inti karakter melalui proses interaksi yang empatik antara konselor (guru bimbingan dan konseling) dengan peserta didik, dimana konselor membantu peserta didik untuk mengenal kelebihan dan kelemahan dalam berbagai aspek perkembangan dirinya, memahami peluang dan tantangan yang ditemukan di lingkungannya, serta mendorong penumbuhan kemandirian peserta didik (konseli) untuk mengambil berbagai keputusan penting dalam perjalanan hidupnya secara bertanggung jawab dan mampu mewujudkan kehidupan yang produktif, sejahtera, bahagia serta peduli terhadap kemaslahatan umat manusia.

Dasar pertimbangan atau pemikiran tentang penyelenggaraan bimbingan dan konseling di Sekolah, bukan semata-mata terletak pada ada atau tidak adanya landasan hukum (perundang-undangan) atau ketentuan dari atas, tetapi yang lebih penting adalah menyangkut upaya memfasilitasi peserta didik agar mampu mengembangkan potensi dirinya atau mencapai tugas-tugas perkembangannya dalam aspek fisik, emosi, intelektual, sosial, dan moral-spiritual.

Di manapun proses pendidikan harus dipandang sebagai suatu proses perkembangan, karena setiap peserta didik adalah seorang individu yang sedang berada dalam proses berkembang atau menjadi, yaitu berkembang ke arah

kematangan atau kemandirian. Untuk mencapai kematangan tersebut, peserta didik memerlukan bimbingan, agar memiliki pemahaman yang baik tentang dirinya dan lingkungannya serta pengalaman dalam menentukan arah kehidupannya.

Alasan lain adalah adanya perbedaan individual pada peserta didik dan keniscayaan bahwa proses perkembangan peserta didik tidak selalu berlangsung secara mulus, dalam alur yang lurus, searah dengan potensi, harapan dan nilai-nilai yang dianut.

Upaya dalam menanggulangi masalah tersebut adalah dengan mengembangkan potensi peserta didik dan memfasilitasi mereka secara sistematik, terprogram dan kolaboratif untuk mampu mencapai standar kompetensi nilai perkembangan/perilaku yang diharapkan. Upaya ini merupakan wilayah garapan bimbingan dan konseling yang harus dilakukan secara proaktif, intensional dan kolaboratif yang diselenggarakan dengan berbasis data perkembangan peserta didik secara komprehensif dalam berbagai aspek kehidupannya.

Atas dasar itu, maka implementasi bimbingan dan konseling di Sekolah diorientasikan kepada upaya memfasilitasi perkembangan potensi peserta didik, yang meliputi aspek pribadi, sosial, belajar, dan karir; atau terkait dengan pengembangan pribadi peserta didik sebagai makhluk yang berdimensi biopsikososiospiritual (biologis, psikis, sosial, dan spiritual).

Salah satu potensi yang dimiliki individu sebagai garapan dari bimbingan dan konseling adalah intelegensi. Intelegensi merupakan modal dasar dari setiap individu untuk mencapai perkembangan optimal yang diharapkan. Hal ini sesuai dengan pernyataan Weschler (dalam Jackson, 2003) yang mengatakan bahwa inteligensi merupakan kapasitas keseluruhan atau global individu untuk bertindak, berpikir rasional, dan menangani lingkungan secara efektif. Artinya bahwa dalam proses perkembangannya, setiap individu akan mengalami keragaman hal ini didasari oleh kapasitas intelegensi yang berbeda pada setiap orangnya.

Dalyono (Djamarah, 2002:160) yang menyebutkan secara tegas bahwa "seorang yang memiliki intelegensi baik (IQ-nya tinggi) umumnya mudah dalam belajar dan hasilnya cenderung baik, sebaliknya orang yang intelegensinya rendah, cenderung mengalami kesukaran dalam belajar, lambat berfikir, dan prestasi yang rendah". Pernyataan serupa dikatakan oleh Cahaya Prabu (2002:161) yang menyatakan bahwa jika siswa yang memiliki tingkat intelegensi tinggi berada dalam lingkungan yang menunjang, maka mereka akan dapat mencapai prestasi dan keberhasilan dalam hidupnya.

Sebagai layanan yang profesional maka layanan bimbingan dan konseling harus memperhatikan kebutuhan siswa. William J. Kolarik (Nurihsan, 2006: 55) mengungkapkan bahwa kualitas mutu layanan bimbingan akan mendapatkan pengakuan jika layanan bimbingan dan konseling mampu memenuhi apa yang diharapkan oleh para konseli. Berdasarkan paparan diatas jelas kiranya bahwa intelegensi sebagai bagian terpenting dari proses perkembangan individu, merupakan hal penting yang harus dikaji lebih mendalam dalam bimbingan dan konseling sebagai upaya memenuhi hakikat tugasnya sebagai komponen penting dalam proses pendidikan.

2. Pemahaman Tentang Inteligensi dalam Konteks Bimbingan dan Konseling

Pendidikan merupakan upaya membantu peserta didik mencapai kedewasaan. Salah satu langkah kegiatan yang sangat menentukan keefektifan pencapaian tujuan pendidikan adalah pemahaman tentang karakteristik anak. Pemahaman tentang anak sangat penting dalam menentukan strategi pendidikan, penentuan tujuan suatu kegiatan yang mungkin dicapai, menentukan metode yang tepat, demikian juga dalam menentukan bantuan khusus yang diperlukan oleh anak tersebut.

Bimbingan dan konseling yang merupakan bagian integral dari pendidikan, efektivitasnya juga banyak ditentukan oleh pemahaman pembimbing tentang diri anak atau klien; bahkan dikatakan, pemahaman terhadap anak

merupakan bagian integral dari proses bimbingan dan konseling itu (Hood and Johnson, dalam Nurhudaya, 1999: 28). Upaya memahami anak dalam proses bimbingan tidak hanya dilakukan pada awal atau sebelum bimbingan, tetapi juga selama proses bimbingan.

Bimbingan bertujuan membantu siswa mencapai perkembangan yang optimal, yang antara lain dengan membantu anak membina keterampilan memahami diri, memahami lingkungan, mengarahkan diri, mengambil keputusan, yang diharapkan dapat mengantarkan anak untuk secara mandiri mengaktualisasikan dirinya. Untuk itu, pemahaman tentang karakteristik anak, baik kelebihan maupun kelemahannya perlu dilakukan oleh pembimbing.

Salah satu karakteristik yang perlu dipahami oleh pembimbing khususnya, adalah potensi intelektual anak atau salah satunya yang populer dengan istilah inteligensi. Meskipun bukan merupakan faktor dominan dalam menentukan keberhasilan studi anak, inteligensi banyak berpengaruh terhadap keberhasilan belajar anak di sekolah dan juga dapat dijadikan prediksi tingkat pendidikan yang dapat dicapai oleh anak. Selain itu, inteligensi juga diakui berpengaruh terhadap tingkat pencapaian karir atau prestasi kerja seseorang.

Dewasa ini pemahaman terhadap inteligensi tidak terbatas hanya pada inteligensi intelektual saja (kognitif) tetapi juga inteligensi nonintelektif, baik inteligensi sosial, inteligensi emosional, inteligensi praktikal atau inteligensi lainnya, yang dalam aktualisasinya satu sama lain saling melengkapi atau satu sama lain memiliki peran- peran tertentu dalam perilaku individu.

3. Peran Tes Intelegensi dalam Layanan Bimbingan dan Konseling

Secara umum, tes psikologis digunakan untuk mendapatkan informasi. Informasi yang dihasilkan dapat digunakan untuk seleksi, penempatan, diagnosis, evaluasi, membuat perencanaan atau hanya 'sekedar' pemahaman diri atau meningkatkan minat mengikuti konseling. Tes psikologis dapat digunakan, tidak saja untuk mengkonfirmasikan apa yang telah diketahui guru kelas dan orang tua tentang teste, tetapi juga sebagai bagian dari pengambilan keputusan praktis.

Contoh nyata dari pernyataan tersebut adalah ketika dua orang siswa menampilkan perilaku yang sama di sekolah (kelas) atau di rumah, belum tentu memiliki latar belakang atau motif yang sama. Tes psikologis dapat membantu menembus apa di balik perilaku yang nampak dan memberikan arah untuk suatu keputusan yang bijaksana atas dasar pemahaman yang mendalam tentang siswa beserta motivasi-motivasinya. Dalam hal inilah tes psikologis dapat memberikan manfaat yang sebaik-baiknya dalam setting sekolah.

Informasi yang diperoleh dari hasil tes psikologis dapat dijadikan masukan yang sangat berharga bagi pembimbing untuk menentukan strategi apa yang paling baik dalam membantu siswa, membantu pembimbing dalam menggambarkan perilaku anak dan mengidentifikasi tipe lingkungan belajar yang baik bagi anak atau untuk menempatkan anak pada lingkungan belajar yang sesuai.

Secara singkat dapat dikemukakan bahwa tujuan penggunaan tes psikologis dalam bimbingan konseling antara lain : 1). Penempatan dan seleksi, 2). Pemahaman tentang gambaran anak, kekuatan dan kelemahannya, 3). Memperkirakan keberhasilan studi atau latihan tertentu, dan 4). Membantu konselor mengevaluasi keberhasilan proses bimbingannya. Secara rinci Goldman (Subino 1984 : 9) menyebutkan penggunaan tes di dalam pekerjaan bimbingan dan konseling adalah sebagai alat pengumpul keterangan-keterangan untuk keperluan diagnosis, pedoman proses bimbingan dan konseling selanjutnya, untuk membantu percakapan antara konselor dengan konseli, dan untuk membantu konseli mengambil keputusan sendiri apa yang akan ia kerjakan setelah proses bimbingan dan konseling.

Tes psikologis bukanlah satu-satunya cara untuk mencapai tujuan-tujuan di atas, juga bukan merupakan syarat untuk memasuki konseling; tes psikologis hanya merupakan prosedur atau instrumen, dan seperti halnya penggunaan instrumen lain, sering ditemukan kesalahan dalam penggunaan tes psikologis kesalahan penggunaan tes antara lain meliputi: 1) Penggunaan skor tunggal untuk menggambarkan perilaku atau penampilan anak, 2) Penggunaan skor atau data

tunggal yang diambil pada satu periode untuk digunakan terus menerus, 3) Memfokuskan hanya pada skor tunggal tanpa melihat data lain, dan 4) Terikat dengan kuantifikasi.

Selain itu beberapa kesalahan yang mungkin dilakukan konselor dalam penggunaan tes antara lain : 1) Menggunakan tes untuk suatu tujuan di luar tujuan pengembangan tes, 2) Menggunakan tes pada kelompok yang berbeda dengan kelompok standarisasi tes, 3) Menganggap tes menjadi 'adil budaya' dengan menerjemahkannya ke dalam bahasa anak, 4) Menganggap tes yang dikembangkan pada sampel yang besar juga dapat digunakan bagi anak yang memiliki handikap tertentu, dan 5) Gagal menerjemahkan hasil tes untuk dapat dipahami oleh testi atau orang tuanya. (Muro & Kottman, 1995 : 337-339).

Beberapa kritik terhadap tes umumnya ditujukan pada penggunaan tes dan bukan pada materi atau bentuk tes itu sendiri. Tes Psikologis (khususnya tes Inteligensi) merupakan suatu instrumen atau alat yang baik buruknya sangat ditentukan oleh penggunanya. Bila dipergunakan sesuai dengan tujuan pengembangannya tes dapat membantu penggunanya untuk mengambil keputusan dengan tepat dan efisien. Sebaliknya bila tidak tepat penggunaannya, selain akan membuat keputusan yang tidak tepat, juga dapat menimbulkan kesalahan perlakuan terhadap anak dan juga membuang waktu dan biaya.

Tes hanya menghasilkan sebagian informasi tentang anak atau lebih tepat lagi hanya mengungkap sampel perilaku anak. Oleh karena itu, upaya pemahaman terhadap anak tidak cukup hanya bertumpu pada hasil tes (skor tes/angka) tanpa memahami lingkungan tempat anak berada/tinggal, dan sistem yang berlaku di lingkungan tersebut.

Beberapa hal perlu dipertimbangkan sebelum penggunaan suatu tes baik untuk konseling, pendidikan maupun untuk keperluan lain, antara lain :

a. Pemahaman atau keahlian konselor atau pembimbing dalam penggunaan suatu tes. Bila konselor atau pembimbing tidak memiliki pengetahuan atau keterampilan melaksanakan tes psikologis, sebaiknya meminta bantuan pihak yang kompeten atau menggunakan cara atau instrumen lain.

b. Pemahaman pihak pengguna hasil tes tentang makna tes psikologis dan data yang dihasilkannya. Pihak pengguna jasa layanan tes perlu mengetahui apa tes psikologis itu dan data seperti apa yang dihasilkannya. Hal ini dimaksudkan agar tidak terjadi kesalahan 'perlakuan' terhadap tes dan hasilnya.
c. Persetujuan dari individu yang akan dites untuk melaksanakan tes psikologis. Hal ini perlu diketahui karena ketepatan hasil tes psikologis akan sangat tergantung pula pada bagaimana individu mengaktualisasikan diri atau bagaimana individu tersebut merespon tes.
d. Tujuan yang ingin dicapai dengan penggunaan tes. Tes psikologis umumnya dibuat berlandaskan pada konsep tertentu dan mengungkap aspek tertentu. Tes psikologis tidak tepat digunakaan untuk tujuan di luar jangkauan tes tersebut.
e. Kualitas tes psikologis. Tes psikologis yang baik teruji secara empirik, baik reliabilitas, validitas, norma yang digunakan untuk penafsiran atau persyaratan yang lainnya seperti kemudahan skoring, tingkat kesulitan, daya pembeda, dan sebagainya.

4. Manfaat Tes MAB – II untuk Bimbingan dan Konseling

Secara umum yang didapat dari hasil tes MAB II sebagai instrumen pengungkap inteligensi siswa, skor yang diperoleh dapat membantu pembimbing untuk lebih mengungkap permasalahan yang dihadapi anak dan membantu pembimbing dalam menentukan strategi dan materi bimbingannya.

Sebagai contoh, Anak yang skor subtes *Information* dan *Comprehension* nya sangat rendah diperkirakan memerlukan dorongan untuk meningkatkan minat belajar dan bimbingan sosial-pribadinya. Implikasi-implikasi dari pencapaian skor siswa dalam setiap subtes MAB II seperti terurai pada bagian deskripsi subtes, dapat dijadikan acuan atau rambu-rambu baik untuk menentukan strategi bimbingan maupun dalam menentukan materi bimbingan. Selain itu respon-respon anak selama tes berlangsung juga sering merupakan informasi yang lebih penting dibanding jawaban anak atas pertanyaan soal tes

dan sangat berharga untuk diperhatikan dan 'ditindak lanjuti' oleh tester, apalagi bila yang menjadi tester adalah pembimbingnya.

Manfaat lain dari penggunaan MAB II dalam bimbingan dan konseling secara lebih khusus adalah sebagai berikut:

a. MAB- II merupakan suatu instrumen pengukuran inteligensi siswa yang memiliki rentang usia sasaran cukup lebar, yakni dimulai dari usia 16 tahun hal ini memungkinkan penggunaan MAB II untuk anak usia dan SMA dan sederajat sebagai landasan pengembangan program bimbingan dan konseling di sekolah.
b. MAB – II memiliki 2 sub tes yaitu verbal dan performa yang masing masing sub tes terdiri atas 5 bagian yang secara kesatuan mengungkap kemampuan global individu dan mampu memberikan rambu-rambu yang lebih jelas tentang potensi yang dimiliki siswa khususya dalam hal intelegensi dan bakat. Bagi kegiatan bimbingan dan konseling, hal ini merupakan hal yang sangat penting untuk difahami dan dikembangkan dalam rangka mengopimalkan potensi siswa di sekolah.
c. Penelitian ini juga diharapkan dapat memberikan sumbangan bagi pengembangan dan adaptasi tes intelegensi khususnya pada tes MAB II sebagai upaya pemahaman lebih lanjut tentang kondisi intelegensi siswa dalam konteks bimbingan dan konseling.

B. Inteligensi

1. Teori – Teori Inteligensi

a. Teori Klasik tentang Inteligensi

Di dalam upaya mendefinisikan hakekat inteligensi terdapat definisi yang berbeda-beda. Perbedaan-perbedaan terserut lazimnya bersumber pada perbedaan-perbedaan pengertian dasar yang dibangun dari renungan tentang dan atau pengamatan terhadap apa yang disebut dengan prilaku *intelligent* itu. Teori- teori yang kemudian menjadi sumber perbedaan dari definisi-definisi tersebut disajikan sebagai berikut,

1) Teori Keturunan – Lingkungan

Teori ini mempunyai tiga anak teori. Anak teori pertama adalah yang memandang inteligensi lebih ditentukan oleh keturunan daripada oleh lingkungan. Salah seorang tokoh dari anak teori ini adalah Arthur H. Jensen. Dengan studinya, baik dengan menganalisis faktor kebersamaan (*common factor*) dari berbagai tes inteligensi yang dikatakannya sebagai yang menuntut proses mental yang tinggi maupun menunjuk kepada program pendidikan yang gagal meningkatkan IQ anak-anak, dan menunjuk kepada perbedaan ras dan status sosial yang terbukti tidak mampu menjelaskan terjadinya perbedaan inteligensi antara orang yang satu dengan orang yang lain, Jensen (Jakson, 2003) berkesimpulan bahwa inteligensi itu lebih ditentukan oleh keturunan dari pada oleh lingkungan. Menurut Jensen, yang merupakan sumber perbedaan antara individu yang satu dengan individu yang lain adalah faktor *"g"* .

Tokoh kedua yang cukup terkenal di dalam anak teori ini adalah Sir Cyril Burt. Di dalam konsepsinya Burt memandang inteligensi itu sebagai kemampuan berfikir umum yang dibawa sejak lahir. "Burt (1955*) is almost alone in grasping this nettle firmly and defining intelligence as innate general cognitive ability*" (Butcher, 1973 : 28). Burt selanjutnya memandang inteligensi itu sebagai sarana potensial untuk meningkatkan penyamaan sosial dan menjamin bahwa anak-anak yang cerdas, sekalipun terbenam di dalam kemelaratan dan hambatan-hambatan lingkungan, namun akan memiliki kesempatan untuk meraih peindidikan mereka dengan potensi mereka.

Di dalam penelitiannya terhadap populasi yang bertempat tinggal di kota London pada tahun 1950-an, Burt menemukan bahwa urunan faktor keturunan terhadap IQ, di dalam artian sempit, berkisar antara 45% sampai dengan 60%. Sedangkan di dalam arti luas, urunan tersebut berkisar antara 80% sampai dengan 85% .

Tokoh ketiga, yang sangat terkenal karena usahanya mengembangkan tes inteligensi individual, yang juga memandang inteligensi sebagai kapasitas bawaan, adalah David Wechsler. Ia memandang inteligensi itu sebagai

kapasitas yang bulat untuk bertindak secara terarah, berfikir rasional, dan berhubungan dengan lingkungannya secara efektif Wechsler (Jakson 2003: 45) mengatakan bahwa "*Intelligence is the aggregate or global capacity of the individual to act purposefully, to think rationally and to deal effectively with his environment*".

Anak teori yang kedua adalah yang memandang inteligensi sebagai yang lebih ditentukan oleh lingkungan daripada oleh keturunan. Salah seorang tokohnya adalah Jerome S. Kegan. Di dalam pengamatannya terhadap anak-anak kulit putih lapisan bawah dan lapisan menengah, sewaktu mereka mengerjakan tes inteligensi, Kegan melihat bahwa anak-anak lapisan bawah bekerja kurang baik apabila dibandingkan dengan anak-anak lapisan menengah. Menurut dia, hal itu disebabkan karena anak-anak lapisan bawah tidak mengalami mesranya hubungan dengan orang tua mereka seperti yang dialami oleh anak-anak lapisan menengah, sebagai contoh oleh Kegan disebutkan bahwa anak-anak lapisan bawah hampir tak pernah tatap wicara dengan orang tua mereka, orang tua mereka hampir tak pernah senyum kepada mereka, dan juga hampir tak pernah bercanda. Hal yang demikian, menurut Kegan, menghambat perkembangan mental anak sehingga menyebabkan rendahnya skor tes inteligensi yang dapat mereka capai (Kegan, 1969).

Anak teori yang ketiga adalah yang memandang inteligensi sebagai hasil antara keturunan, lingkungan dan interaksi antara keduanya. Jadi inteligensi setiap orang itu selalu mendapat urunan dari keturunan, lingkungan dan interaksi antara keduanya.

Tokoh yang mempunyai konsepsi ke tiga ini adalah adalah Alan O. Ross. Ia menyatakan bahwa "*measured intelligence is a complex function of the interaction of many genetic ana environmental factors*". Oleh sebab itu secara teoretis konsepsi ini akan menyatakan bahwa tingkat pendidikan yang mungkin dapat dicapai oleh seseorang harus dikembalikan kepada inteligensi yang dimilikinya, lingkungan tempat ia tinggal, dan interaksi keduanya. Pengertian

interaksi antara inteligensi dengan lingkungan di sini adalah dalam arti penggunaan lingkungan untuk mengembangkan inteligensi yang bersangkutan.

2) Teori Epistemologis – Biologis

Anak teori pertama dari teori ini memandang inteligensi sebagai kemampuan berfikir jernih, analitis, dan komprehensif. Tokoh pertama yang termasuk ke dalam kelompok anak teori ini adalah Lewis M. Terman. Oleh Butcner, pandangan Terman tentang inteligensi itu disarikan sebagai kemampuan berfikir abstrak.

Tokoh ke dua dari anak teori ini adalah G. D. Stoddard. Menurut konsepsinya inteligensi itu merupakan kemampuan yang majemuk, yakni kemampuan menyelesaikan tugas-tugas yang sulit, rumit, abstrak, ekonomis, adaptif terhadap tujuan, berbobot sosial dan original, serta tetap memelihara kemampuan menyelesaikan tugas-tugas seperti itu di dalam keadaan yang menuntut pemusatan energi dan menahan gejolak-gejolak emosional.

Tokoh ke tiga, yang juga memandang inteligensi sebagai kemampuan yang majemuk, adalah Henry E. Garrett. Menurut dia, inteligensi itu paling sedikit sebagai kemampuan-kemampuan yang diperlukan untuk memecahkan persoalan-persoalan yang menuntut pemahaman dan penggunaan simbol-simbol, baik berupa kata-kata, angka-angka, diagram-diagram, persamaan-persamaan maupun rumus- rumus yang menyatakan gagasan-gagasan dan hubungan berbagai hal, dari yang sederhana sampai dengan yang sangat rumit.

Anak teori yang ke dua adalah yang .nemandang inteligensi sebagai kemampuan menyesuaikan diri terhadap situasi yang baru (biologis). Salah seorang tokohnya, yang kini mulai terkenal di Indonesia, adalah Jean Piaget. Menurut Piaget, inteligensi adalah kemampuan melakukan penyesuaian diri terhadap lingkungan. Inteligensi itu oleh Piaget tidak dipandang sebagai sesuatu kekuatan yang berdiri sendiri; ia selalu berkaitan dengan lingkungannya. Kata Piaget (Jakson 2003: 40) , "*Intelligence does not therefore appear as a power of*

reflection independent of the particular position which the organism occupies in the universe but is linked, from the very outset, by biological apriorities"

Menurut Piaget lebih lanjut, inteligensi itu tidak muncul sekaligus di dalam perkembangan mental melainkan muncul sebagai satu kesinambungan yang berlanjut. Atas dasar teori inilah Piaget menyusun tahap-tahap perkembangan inteligensinya dari yang kongkrit sampai dengan yang formal. Pemunculan kemampuan berfikir dari tahap menggunakan reflex sampai dengan tahap deduksi tersebut oleh Piaget dipandang sebagai tahap-tahap perkembangan yang bulat.

Tokoh ke dua dari kelompok anak teori ke dua ini adalah William Stern, Ia berpendapat bahwa inteligensi itu adalah kemampuan personal untuk dapat menghadapi tuntutan-tuntutan baru dengan menggunakan alat-alat berfikir secara efisien.

3) Teori Faktorial

Teori faktorial ini mempunyai berbagai variasi yang masing-masing akan dibahas sebagai berikut.

a) Teori satu faktor Binet

Teori ini berpendapat bahwa inteligensi itu terbangun atas satu faktor saja. dengan lebih singkatnya, inteligensi adalah kesanggupan belajar. Bahwa skala inteligensi Stanford-Binet ini hanya mengukur faktor "g" saja.

b) Teori Dua faktor Spearman

Teori ini berpendapat bahwa inteligensi itu terbangun atas-dua faktor. Pelopor teori ini adalah Charles Spearman yang berpendapat bahwa inteligensi itu terbangun atas faktor "g" dan faktor "s". Faktor "g" merupakan kemampuan umum yang bersifat universal, sedangkan faktor "s" merupakan faktor khusus yang khas (Cronbach, Jackson 2003: 47). Menurut Spearman, setiap orang memiliki kedua faktor tersebut. Yang membedakan antara individu yang satu dengan yang lain adalah perbedaan perpaduan keduanya, yang akan merupakan bakat yang terpancarkan, "*manifested aptitude*".

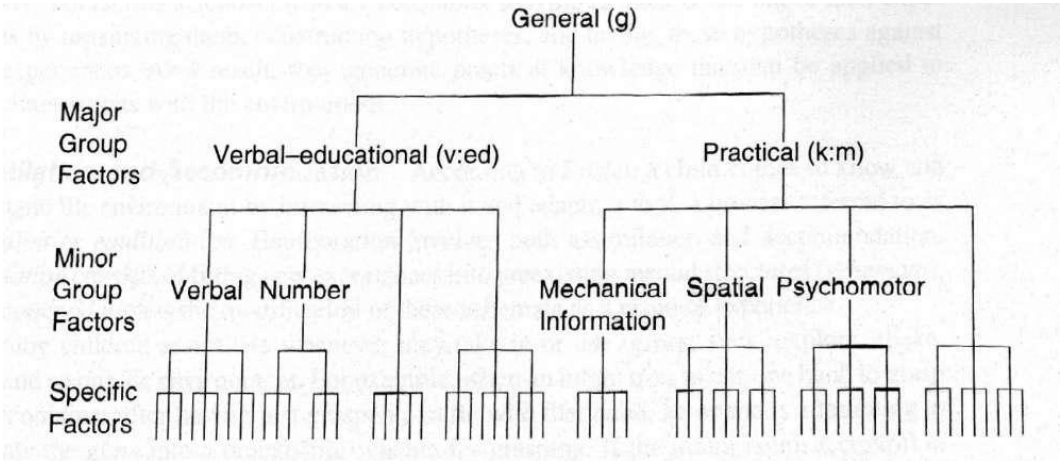
Arti penting teori dua faktor Spearman ini, untuk keperluan analisis tes, ialah asumsinya yang menyatakan bahwa kalau antara dua buah tes terjadi korelasi, maka kedua tes tersebut mengukur satu faktor kebersamaan yang sama dan dua faktor khusus yang berlainan (Guilford, dalam Subino 1986 : 472). Dengan demikian maka kedudukan skor pada tes yang satu dapat dipergunakan untuk meramalkan kedudukan skor tes yang lain. Kesimpulan yang demikian dilandasi oleh pemikiran bahwa kedua tes tersebut menuntut satu kemampuan yang sama.

c) Teori faktor bertingkat Philip E. Vernon

Konsepsi inteligensi Vernon ini mirip dengan konsepsi Spearman. Menurut Vernon, di bawah faktor "g" itu terdapat dua *major group factors* yang masing-masing adalah *verbal - educational* (V : ed) dan *practical* (k : a). Yang pertama dibagi ke dalam dua *minor-group factors*, yakni *verbal* dan *numerical*; sedangkan yang kedua dibagi menjadi *snace ability*, *manual ability*, dan *mechanical ability*. Dan masing-masing bagian tersebut dibagi lagi menjadi *specific factor* vang sangat besar jumlahnya dan mencakup lingkup yang sangat khusus (Vernon, Jakson 2003 : 22). hodel bangunan inteligensi Vernon ini dapat dilukiskan ke dalam bagan berikut.

Bagan 2.1

Teori faktor bertingkat Philip E. Vernon



Dalam model hirarki kemampuan mental Vernon, semakin tinggi posisi faktor dalam diagram, semakin luas rentang perilakunya. Model hirarki Vernon mempertahankan faktor umum inteligensi Spearman dan meletakan *primary mental abilities* Thurstone dan *structure-of-lntellect* Guilford pada status bawah di bawah faktor "g". Model hirarki Vernon adalah cara yang baik dalam menggabungkan temuan dan interpretasi berbagai studi analisis faktor menjadi sebuah teori tunggal. Konsep yang dikemukakan Vernon serupa dengan Wechsler.

Weschler merancang tes inteligensi yang terdiri dari dua bagian yaitu *verbal* dan *performance* yang pada akhirnya digabungkan menjadi skor IQ.

4) Teori Struktural

a) Teori struktural Guilford

Dalam teori yang pada dasarnya merupakan pengembangan dari teori *Primary Mental Abilities* Thurstone, Guilford mengusulkan bahwa kinerja pada setiap tugas kognitif terbaik dapat dipahami melalui analisis ke dalam jenis operasi mental atau proses mental yang dilakukan, jenis konten atau uji materi dimana operasi mental dilakukan, dan produk yang dihasilkan dari operasi tertentu pada jenis uji konten tertentu. Model *Structare-of-lntellect* berisi 120 faktor yang berbeda, yang dianggap independen. Pengembangan dari teori ini mengungkapkan bahwa inteligensi asli terdiri dari lima macam jenis operasi (evaluasi, produksi konvergen, produksi divergen, pemahaman memori, perekaman memori, kognisi), enam jenis produk (unit, kelas, hubungan, sistem, transformasi, dan implikasi), dan lima jenis isi (visual, pendengaran, simbolik, semantik, perilaku). Hal ini menunjukkan adanya 150 (5 X 5 X 6) tugas intelektual yang terdiri dari struktur intelek.

Walaupun awalnya Guilford berasumsi bahwa 150 faktor tersebut independen satu sama lain, tetapi tidak ada penelitianya yang mendukung asumsi ini. Guilford kemudian mengganti model *Structure of lntellect* yang telah dimodifikasi dengan model kemampuan hirarkis yang terdiri dari 150 orde faktor pertama, 85 orde faktor kedua, dan 16 orde faktor ketiga. Model terakhir belum sepenuhnya dievaluasi oleh penelitian, tapi Brody (dalam Aiken, 1997)

menyimpulkan bahwa itu bukan merupakan alternatif yang dapat diterima untuk model hirarkis yang mencakup faktor umum “g” di puncak.

b) Teori Faset Guttman

Teori model PASS (*Flanning - Attention Ünndtaneous Processing - Successive Processing)* menyatakan bahwa aktivitas kognitif otak manusia dibagi menjadi tiga unit fungsional. Unit fungsional pertama, yang dianggap berhubungan dengan batang otak bagian atas dan sistem limbik, berhubungan dengan gairah atau perhatian dan adalah pembeda antara rangsangan. Walaupun tidak bertanggung jawab untuk menerima dan menganalisis informasi, unit ini sangat penting untuk proses kognitif karena memberikan kesiapan kondisi umum dan fokus perhatian.

Secara ringkas, unit fungsional pertama dikatakan bertanggung jawab atas gairah dan perhatian, yang kedua untuk penerimaan, analisis, dan penyimpanan menggunakan proses penalaran secara berturut-turut dan simultan, dan ketiga untuk merencanakan, mengatur, dan memverifikasi aktivitas mental. Baik input dan output mungkin serial atau bersamaan.

b. Teori Modern tentang Inteligensi

Beberapa tahun terakhir ini, terjadi peningkatan minat terhadap penelitian terkait inteligensi yang menghasilkan sejumlah konsepsi baru, perhatian terhadap aspek baru dari proses-proses tertentu, dan pemahaman yang semakin dalam terkait isu seputar inteligensi (Sternberg, dalam Anastasi 1997: 46).

Teori Guilford membahas inteligensi dalam kaitannya dengan tiga dimensi yang luas yakni operasi, konten dan produk. Dimensi operasi mencakup kognisi, memori, produksi divergen, produksi konvergen, dan evaluasi. Dimensi konten mencakup figural, simbol, semantik, dan behavioral. Dimensi produk mencakup unit, kelas, relasi, system, transformasi, dan implikasi. Ketiga dimensi tersebut membentuk matriks dengan 120 kotak. Fokus penelitian Guilford adalah untuk mengidentifikasi faktor-faktor yang berbeda untuk mengisi 120 kotak yang

tersedia dalam matriks. Kemungkinan apakah faktor-faktor tersebut saling berkorelasi dan terorganisasi secara hierarkis menjadi pertanyaan yang belum terjawab dari penelitian Guilford tersebut.

Carroll (Jakson:2003) melakukan studi dengan menggunakan teori dasar mengenai pemrosesan kognitif. Studi yang dilakukan Carroll bertujuan untuk mengidentifikasi pemrosesan yang berbeda-beda dalam diri individu yang mungkin berkaitan dengan masing-masing faktor. Dalam sistem yang dibuat Carroll, faktor kemampuan mental yang utama muncul ketika seperangkat tes membagi perbedaan individual menjadi empat ranah yakni jenis dan konten memori (contoh : *leksikosemantik, quantitative*), lingkup dan urutan pengoperasioan kognitif, strategi kognitif dan pemecahan masalah, serta kelas respon. Proses kontrol, menurut Carroll, terbagi menjadi tipe instruksi (perhatian, pengarahan, dan memori jangka panjang) serta tipe pemecahan masalah (pengoperasian terkait fungsi eksekutif atau integratif dari memori jangka pendek).

Sternberg (Anastasi 1997: 45) melakukan penelitian eksperimen modern terkait pemrosesan informasi dan kognitif, seperti halnya penelitian inteligensi yang menggunakan pendekatan tradisional, penelitian tersebut menggunakan 'teori triarkis inteligensi manusia'. Teori triarkis terdiri atas tiga subteori yakni : subteori kontekstual, yang membahas cara bagaimana individu beradaptasi terhadap pemilihan dan pembentukan lingkungan eksternal; subteori dua-segi berkenaan dengan peran pengalaman dalam mempengaruhi reaksi terhadap pembaharuan dan automatisasi; dan subteori yang menghubungkan inteligensi dengan dunia internal individu. Ketiga subteori tersebut mengidentifikasi proses intelektual yang memfasilitasi perilaku intelijen individu. Subteori yang ketiga menggarisbawahi tiga komponen yakni pemrosesan informasi penting untuk mempelajari, merencanakan dan melaksanakan operasi-operasi yang diperlukan. Gagasan Sternberg dapat dikatakan revolusioner karena : (a) Memperluas definisi inteligensi dalam beberapa bidang, contohnya adaptasi okupasional dan inteligensi sosial; (b) Memberikan hipotesis eksplisit mengenai proses-proses

tersembunyi; dan (c) Merupakan sumber yang kaya akan gagasan penelitian selanjutnya.

Teori modern mengenai inteligensi biasanya lebih luas lingkupnya, oleh karena itu dapat memberikan dasar untuk mengevaluasi konteks dimana tes inteligensi diperlukan, menyarankan isi dan tata cara tes, serta memberi petunjuk tentang bagaimana tes inteligensi berfungsi sebagaimana mestinya dan bagaimana fungsi tersebut dapat dikembangkan. Kajian teoretis yang dilakukan Guilford Caroll (1993), dan Sternberg (1984) memberikan orientasi eksplisit untuk memperluas lingkup inteligensi menjadi sistem yang teorganisir dalam cara tertentu. Sistem tersebut memberikan fondasi yang kaya untuk memformulasi hipotesis terkait area baru di bidang pengukuran.

Isu pertama yang dimunculkan oleh teori inteligensi faktor jamak dalam studi yang dilakukan Thurstone (Anastasi 1997: 62) yakni mengenai apakah formulasi teori yang telah dikembangkan dan didiferensiasi oleh Guilford dapat dipraktikkan dengan tepat. Teori tersebut tentunya telah memperluas dan memperdalam pengetahuan mengenai hal-hal seputar inteligensi dan menyajikan penegasan mengenai strategi pemrosesan-informasi, konten, dan persyaratan tugas beserta dampaknya terhadap individu dan perkembangan teori di masa depan (Sternberg, 1997). Namun, dampak teori tersebut terhadap praktisi di lapangan sangatlah terbatas. Kebanyakan praktisi memilih untuk menggunakan konstruk inteligensi umum (Wechsler, 1958), karena dipandang sebagai pengukuran global terhadap kemampuan intelektual yang penting dalam memahami perilaku individu. Meskipun faktor-faktor yang berdasarkan atas teori tentang kemampuan intelektual telah memberikan jaminan terhadap perediksi diferensial, misalnya bakat khusus memprediksi aspek khusus dari performansi kerja, namun jaminan tersebut belum direalisasikan dalam kebanyakan praktik. Brody dan Brody (Jackson 2003: 40) menyeimbangkan penilaian kritis terhadap teori melalui apresiasi terhadap perlunya persediaan data yang berbasis faktor kemampuan tunggal dengan data lainnya. Mereka mempercayai bahwa teori mengenai struktur intelektual yang berdasarkan atas faktor kemampuan tunggal merupakan

representasi yang kurang memadai dari struktur intelektual. Pandangan tersebut terlepas dari konsistensi bukti-bukti mengenai kelebihan faktor “g” sebagai suatu faktor inteligensi.

Brody dan Brody (Jackson 2003: 52) mengemukakan bahwa dasar untuk mengukur faktor tunggal yang kuat dari kemampuan kognitif adalah fakta bahwa testi/individu bukanlah seorang yang jenius dalam mengembangkan strategi berbeda untuk memecahkan permasalahan yang sama, strategi yang mereka gunakan biasanya disesuaikan dengan kemampuan unik yang dimiliki. Argumen tersebut konsisten dengan kedudukan proses eksekutif atau proses metakomponensial Sternberg yang mengalokasikan sumber-sumber intelektual dan mengendalikan strategi-strategi pemrosesan informasi. Proses-proses tersebut tidak sama antara individu satu dengan individu lainnya namun hasilnya mungkin saja sama. Selain itu, penting untuk meringkas keseluruhan performa menjadi satu abstraksi tunggal, yang disebut skor total, dikarenakan oleh tiga alasan : (a) Keragaman memunculkan kesempatan untuk pengalaman yang tidak disengaja dan sumbangan berupa materi yang spesifik, (b) Fungsi-fungsi eksekutif dapat muncul secara diferensial, dan (c) Percampuran tugas yang memungkinkan adanya penemuan baru dan automatisasi dianggap tepat.

c. Intelegensi sebagai Konstruk MAB II

Di dalam upaya mendefinisikan hakekat intelegensi akan terdapat definisi yang berbeda-beda. Perbedaan tersebut lazim bersumber pada perbedaan-perbedaan pengertian dasar yang dibangun dari renungan tentang dan atau pengamatan terhadap apa yang disebut dengan perilaku intelegent. Hal itulah yang mendasari para tokoh dalam mendefinisan intelegensi tersebut. Sejumlah penyusun tes telah berupaya mendefinisikan inteligensi, salah satu nya adalah seperti yang dikemukakan oleh Wechsler (dalam Jackson, 2003):

Inteligensi merupakan kapasitas agregat atau global dari individu untuk bertindak secara bertujuan dan berpikir secara rasional serta berhubungan secara efektif dengan lingkungan sekitarnya. Disebut agregat atau global, karena inteligensi terdiri atas elemen-elemen atau kemampuan-kemampuan yang meskipun tidak sepenuhnya independen tapi dapat dibedakan secara kualitatif.

Wechsler (dalam Jackson, 2003) mengemukakan bahwa perilaku intelijen merepresentasikan lebih dari sekedar penyatuan beragam kemampuan yang terpisah-pisah, karena : (a) Kombinasi dari kemampuan-kemampuan tertentu terbukti lebih manjur dibanding lainnya; (b) Kualitas lain, misalnya motivasi, mempengaruhi perilaku; dan (c) Tingkat kemampuan yang lebih tinggi dari yang dipersyaratkan suatu tugas tidak akan mempengaruhi perilaku tertentu. Terdapat beberapa implikasi dalam diskusi Wechsler mengenai hal-hal seputar inteligensi, diantaranya : (1) Inteligensi mencakup beberapa kemampuan terpisah yang terorganisir secara bersama-sama, (2) Inteligensi dan perilaku intelijen merupakan dua hal yang berbeda, (3) Pengaruh lain diluar kemampuan dapat berdampak terhadap perilaku intelijen, dan (4) Tingginya kemampuan dalam area tertentu tidak harus atau tidak selalu menghasilkan perilaku intelijen di area lainnya.

Gagasan tersebut dan gagasan lain yang relevan memiliki latar belakang sejarah, dan sampai sekarang masih menjadi perhatian para pakar teori kontemporer. Latar belakang sejarah terkait inteligensi memberikan dasar yang kuat terhadap pembahasan isu-isu seputar teori inteligensi, strategi pengukuran, peran konteks budaya, dan kecenderungan di masa depan.

Apabila inteligensi sebagai konstruk hipotetik dipandang berbeda dari apa yang diukur oleh tes inteligensi, maka beberapa teori tentang inteligensi perlu memiliki pengakuan dari sejumlah besar literatur empirik yang memadai dalam hal konseptualisasi. Berikut dipaparkan beberapa disposisi mengenai inteligensi yang ditarik dari pemikiran kontemporer terkait kemampuan intelektual (Sternberg, 1997). Preposisi-preposisi ini sebaiknya tidak dipandang sebagai fakta ataupun suatu pernyataan tanpa kontroversi. Alih-alih preposisi ini dapat

dipandang sebagai hipotesis yang menjadi dasar bagi pengukuran, penelitian, serta klarifikasi dan revisi di masa mendatang.

Berikut adalah preposisi-preposisi mengenai kemampuan intelektual sebagai landasan pengembangan tes MAB II (Jackson, 2003) :

1) Kemampuan inteligensi dan intelektual merepresentasikan konstruk hipotetik yang ditarik dari berbagai studi observasional, psikometrik, laboratoris, dan studi lapangan. Sistem konstruk tersebut tidak dapat diamati secara langsung, namun dapat diketahui melalui pengamatan terhadap seperangkat perilaku yang membentuk konstruk tersebut. Dengan demikian, inteligensi berbeda dari konstruk lain yang digunakan untuk menafsirkan kejadian fisik seperti medan elektromagnetik atau temperature
2) Inteligensi merupakan suatu unitas multiplex, yakni seperangkat sistem yang terdiferensiasi, beroperasi secara semi-otonomi namun tetap menjadi metasistem yang terorganisir. Metasistem tersebut dapat dikonseptualisasikan sebagai suatu aspek kepribadian sehingga perilaku intelijen (contoh : aktivitas yang bertujuan) dapat berfungsi tidak hanya sebagai kemampuan tapi juga sebagai ciri kepribadian misalnya ketekunan dan fleksibilitas.
3) Terdapat ciri khusus dari inteligensi yang muncul seiring perkembangan dikarenakan kematangan fungsi biologis. Inteligensi menurun seiring proses penuaan, sebagian dikarenakan fungsi entropi biologi. Studi yang konsisten dengan preposisi ini adalah studi mengenai hubungan antara masa otak dan inteligensi, serta temuan penelitian yang mengungkap hubungan antara perilaku intelijen dan respon tersembunyi (Vernon dan Weese, 1993). Kecepatan pemrosesan informasi, pemecahan masalah, dan kecepatan reaksi memberikan informasi penting mengenai keberfungsian intelektual dan termasuk aspek yang dipertimbangkan dalam pengukuran kemampuan intelektual.
4) Banyak ragam pemrosesan intelektual yang berpengaruh terhadap perilaku intelijen. Pemrosesan tersebut dapat terdeteksi oleh konten dan format tes. Korelasi antara tes dan kemunculan kelompok faktor akan menjadi suatu

fungsi dimana pemrosesan serupa muncul dalam tugas-tugas intelektual yang berbeda. Sebagian besar pengukuran terhadap kemampuan kognitif saling berhubungan secara positif karenanya perilaku intelijen dapat digambarkan sebagai faktor general, bersama-sama dengan seperangkat faktor intelektual dan non-intelektual.

5) Meskipun salah satu dari banyak ragam performa intelektual dapat digunakan untuk mengukur kemampuan intelektual umum, generalisabilitas akan lebih baik apabila evaluasi dilakukan terhadap sampel yang luas mencakup beragam tipe pengukuran yang berbeda-beda. Semakin representatif sampel, semakin pemrosesan intelektual akan semakin dapat digeneralisasi. Dari sudut pandang teknis, apabila kombinasi skor tes memiliki komponen variansi dan kovariansi, maka kovariansi antara beragam performa yang berbeda akan terakumulasi ketika skor ditambahkan secara bersamaan, menghasilkan perkiraan yang lebih baik atas kesamaan-kesamaan antara seperangkat pengukuran dengan lainnya.
6) Perilaku intelijen merupakan suatu hal yang bergantung konteks dan ditetapkan secara sosial, terdapat kesamaan antara penyesuaian sosial dan perilaku intelijen. Mustahil untuk mengukur perilaku intelijen dalam isolasi terhadap suatu lingkungan dimana seharusnya perilaku intelijen tersebut ditampilkan. Oleh karena itu tidak dianjurkan mengadministrasikan pertanyaan tes menyangkut pengetahuan bahasa Inggris terhadap responden yang tidak mengenal atau tidak memiliki wawasan mengenai bahasa Inggris. Contohnya, Lynn (Jackson 2003: 54) menunjukkan bahwa ketika tes Skala Inteligensi Wechsler diadministrasikan terhadap para siswa di Jepang dengan menggunakan petunjuk dalam bahasa Jepang, para siswa tersebut tidak menemui kesulitan dalam melakukan tugas-tugas dalam tes tersebut. Skor mereka pun secara signifikan lebih tinggi dari skor para siswa di Amerika Utara. Dengan demikian, sarana asesmen perlu disesuaikan dengan budaya respoden (misalnya, bahasa harus dapat dipahami) dan apabila responden memiliki keterampilan tertentu yang diperlukan untuk mengikuti asesmen

(misalnya membaca), maka level keberfungsian intelektual responden tersebut dapat dinilai dengan baik. Tentu saja lingkungan yang kaya, dimana keterampilan intelektual dapat dipraktikkan dan dimana ada banyak kesempatan untuk mengatasi masalah baru, diharapkan untuk dapat berkontribusi terhadap perkembangan kemampuan intelektual.

7) Perilaku intelijen tidak hanya diwujudkan oleh kemampuan yang tercermin dalam performa kognitif tradisional, namun dalam lingkungan sosial juga. Binet dan Thorndike secara eksplisit mengakui adanya peran inteligensi sosial. Sementara para pakar teori modern memasukkan adaptasi terhadap lingkungan behavioral sebagai dimensi utama dari perilaku intelijen. Jackson (1961) telah menunjukkan bahwa pengamat independen secara reliable memberdakan antara inteligensi sosial dan inteligensi akademik dalam diri responden mereka.

d. Tes Intelegensi

1) Konteks Historis dari Tes Inteligensi

Pada awalnya pengukuran inteligensi telah dipraktekkan di negara Cina sebelum dinasti Han, pengukuran inteligensi dilakukan oleh jenderal Cina untuk menguji rakyat sipil yang ingin menjadi anggota legislatif berdasarkan pengetahuan menulis klasik, persoalan administratif dan manajerial.

Pengukuran inteligensi kemudian dilanjutkan sampai pada masa dinasti Han (200 SM- 200 M), namun seleksi ini tidak lagi untuk legislatif saja, tetapi mulai merambah pada bidang militer, perpajakan, pertanian, dan geografi. Meskipun diawali dengan sedikit mencontoh pada seleksi militer Prancis dan Inggris, sistem ujian telah disusun dan berisi aktivitas yang berbeda, seperti tinggal dalam sehari semalam dalam kabin untuk menulis artikel atau puisi, hanya 1 % sampai dengan 7 % yang diizinkan ikut ambil bagian pada ujian tahap kedua yang berakhir dalam tiga hari tiga malam. Menurut Gregory (2007), seleksi ini keras namun dapat memilih orang yang mewakili karakter orang Cina yang

kompleks. Tugas-tugas militer yang berat cukup dapat dilakukan dengan baik oleh para pegawai yang diterima dalam seleksi fisik dan psikologi yang intensif.

Tokoh psikologi yang berperan antara lain adalah Wundt. Beliau merupakan psikolog pertama yang menggunakan laboratorium dengan penelitiannya mengukur kecepatan berpikir. Wundt mengembangkan sebuah alat untuk menilai perbedaan dalam kecepatan berpikir. Tokoh yang tidak kalah pentingnya adalah Alfred Binet. Kontribusi nyata Binet adalah menciptakan tes inteligensi. Binet dibantu oleh Theophile Simon (1904) membuat instrumen pengukur inteligensi dengan skala pengukuran level umum pada soal-soal mengenai kehidupan sehari-hari sehingga tesnya dikenal dengan nama *Tes Binet-Simon.* Perkembangan selanjutnya dua tokoh ini mengembangkan penggunaan tes inteligensi dengan tiga puluh item yang berfungsi mengidentifikasikan kemampuan sekolah anak (Van Ornum, 2008).

Di Amerika, usaha pertama tersebut dimulai oleh tokoh pencetus istilah "tes mental" yaitu James Mckeen Cattell (1860-1944), yang menerbitkan bukunya *Mental and MaMfiremels* d tahun 1890. Buku ini berisi serangkaian tes inteligensi yang terdir atas 10 jenis ukuran. Ke-10 macam ukuran tersebut adalah (Gregory, 2007):

a) *Dynamometer Pressure*, yaitu ukuran kekuatan tangan menekan pegas yang dianggap sebagai indikator aspek psikofisiologis
b) *Rate of movement*, yaitu kecepatan gerak tangan dalam satuan waktu tertentu yang dianggap memiliki komponen mental di dalamnya.
c) *Sensation areas*, yaitu pengukuran jarak terkecil diantara 2 tempat yang terpisah dikulit yang masih dapat dirasakan sebagai 2 titik berbeda.
d) *Peasue caosingpain,* yaitu pengukuran yamg dianggap berguna dalam diaknosis terhadap penyakit saraf dan dalam mempelajari status kesadaran abnormal.
e) *Least noticabele difference in weight,* yaitu pengukuran perbedaan berat yang terkecil yang masih dapat dirasakan seseorang.

f) *Reaction time for sound,* yang mengukur waktu antara pemberian stimulus dengan timbulnya reaksi tercepat.

g) *Time for naming color»,,* yang dimaksudkan sebagai ukuran terhadap proses yang lebih 'mental' daripada waktu-reaksi yang dianggap reflektif.,

h) *Bisection of a 50-cm line,* yang dianggap sebagai suatu ukuran terhadap akurasi *'space judgment'*

i) *Judgment of KMfecond time,* yang dimaksudkan sebagai ukuran akurasi dalam *'time judgment' A* subyek diminta menghitung 10 detik tampa bantuan apapun).

j) *Number of latters repeated upon once hearing,* yang dimaksudkan sebagai ukuran terhadap perhatian dan ingatan (subyek diminta mengulang huruf yang sudah disebutkan lx)

C. Konsep Pengembangan Alat Ukur Psikologis

1. Pengukuran psikologis

Definisi pengukuran berkiblat pada pemikiran Campbell (Guilford, 1954: 5) yang membatasi pengukuran *(measurement*) sebagai *"...the assignment of numerals to objects or events according to rules"*. Pada buku berjudul *Fundamental Statistics in Psychology and Education* Guilford (1978: 21) menambahkan batasan pengukuran sebagai *"...the assignment of numerals or numbers to objects and events according to logically acceptable rules"*. Batasan ini menekankan bahwa proses pengukuran harus berdasarkan sistem yang dapat diterima dan logis. Sistem angka memiliki persyaratan logis itu, sebab angka memberikan kemungkinan-kemungkinan yang majemuk dalam memanipulasi dengan logika. Jika seorang ahli menerapkan angka dalam menerangkan objek atau peristiwa, maka dengan sendirinya orang itu berkesempatan untuk memperlakukan angka-angka tersebut dalam berbagai cara sistematis untuk mengambil kesimpulan dari apa yang sedang diukurnya.

Pada bagian lain, secara teknis Wiersma & Jurs (1990: 8) mendefinisikan istilah pengukuran sebagai “the *assignment of numerals to objects or events*

that give numerals quantitative meaning". Lain lagi dengan Gronlund (1985: 5) yang membatasi istilah pengukuran sebagai *"... the process of obtaining a numerical description of the degree to which an individual possesses a particular characteristic - by answers the question of "How much?"*. Beberapa ahli lain, yang juga berkecimpung dalam pengukuran, seperti Weitzenhoffer (1951) menyebutkan pengukuran sebagai suatu operasi yang dilakukan terhadap alam fisik oleh pengamat, sedangkan Stevens (1946) mengemukakan pengukuran sebagai pemberian angka atas objek atau kejadian sesuai dengan aturan. Akhirnya, Lord & Novick dan Torgerson (Noer, 1987: 1) memperbaiki definisi Stevens dengan menyatakan bahwa pengukuran lebih dikenakan pada sifat objek daripada terhadap objek itu sendiri. Misalnya, pada saat seorang ahli kimia fisika mengukur masa molekul persenyawaan atau seorang ahli biologi menentukan jumlah bakteri di dalam setetes air kolam, semua ini merupakan pengukuran atas sifat khas dari objek yang diteliti. Dalam pengertian yang serupa, ahli psikologi di sekolah tidak mengukur anak, akan tetapi sifat fisik yang khas dari anak seperti tinggi atau beratnya; atau sifat psikis seperti peningkatan penguasaan kosa kata, kematangan sosial; atau pengetahuan tentang dasar-dasar penjumlahan.

Definisi-definisi tentang pengukuran yang dikemukakan para ahli tersebut sangat beragam, namun Nunnally (1978: 3) mengungkapkan bahwa intinya *"...measurement consist of rules for assigning numbers to objects in such a way as to represent quantities of attributes"*.

Selanjutnya, secara sederhana psikologi dapat didefinisikan sebagai studi tentang tingkah laku, dan sifat atau atribut yang mencirikan tingkah laku individu di lingkungan (rumah, pekerjaan, sekolah atau di masyarakat). Hal ini pada umumnya disebut atribut psikologis (*psychological traits)* (Noer, 1987:2-3).

Dengan demikian, merujuk dari beberapa definisi di atas berkaitan dengan pengukuran serta atribut psikologis dalam ranah psikologi, maka dapat dikatakan bahwa pengukuran psikologis merupakan kuantifikasi atribut-atribut psikologis dengan menggunakan aturan-aturan tertentu yang logis dan dapat diterima secara menyeluruh.

Penerapan aturan-aturan di atas dalam pengukuran secara langsung Sehubungan dengan kuantifikasi. Suryabrata (1999: 3) mengungkapkan bahwa dengan pendekatan kuantitatif ini dapat diperoleh beberapa keuntungan, seperti misalnya: 1) Mungkinnya dilakukan pencatatan secara eksak; 2) Perlunya peneliti mengikuti tata pikir dan tata kerja yang pasti dan konsisten; 3) Memungkinkan peneliti meringkas data dalam cara dan bentuk lebih bermakna dan mudah dianalisis; 4) Mungkinnya pengunaan teknik analisis statistik dan matematis yang dapat diandalkan dalam penelitian ilmiah; dan 5) Tingginya komunikabilitas hasil yang diperoleh.

2. Pengembangan Alat Ukur

Secara ringkas ada dua cara dalam mengembangkan alat ukur (instrumen) yaitu dengan mengembangkan sendiri dan dengan cara menyadur. Natawidjaja (Helma, 2001: 51) mengatakan bahwa dalam mengukur suatu variabel penelitian, seorang peneliti dapat menyusun sendiri instrumen penelitiannya, akan tetapi dalam hal-hal tertentu peneliti dapat menggunakan instrumen yang sudah ada, baik itu yang telah digunakan dalam penelitian sebelumnya maupun berupa instrumen baku dalam bahasa asing.

Bila peneliti mengembangkan instrumen sendiri, ada beberapa langkah yang akan ditempuh yaitu: 1) Merumuskan masalah penelitian; 2) Merumuskan variabel; 3) Menentukan instrumen yang akan digunakan; 4) Menjabarkan bangun setiap variabel; 5) menyusun kisi-kisi; 6) Penulisan butir-butir instrumen 7) mengkaji ulang instrumen tersebut yang dilakukan peneliti sendiri atau oleh penilai (melakukan *judgement*); 8) Penyusunan perangkat instrumen sementara; 9) Melakukan uji coba dengan tujuan untuk mengetahui; 10) Perbaikan instrumen setelah hasil uji coba; dan 11) Penataan kembali perangkat instrumen yang

terpakai untuk memperoleh data yang akan digunakan untuk keperluan pembuatan manual.

Bila peneliti mengembangkan instrumen dengan prosedur adaptasi (menyadur) langkah-langkah berikut dapat dilakukan: 1) Menterjemahkan butir-butir pernyataan oleh dua orang penterjemah yang terpisah ke dalam bahasa peneliti (contoh bahasa Indonesia); 2) Peneliti menyunting dan mengintegrasikan hasil terjemahan; 3) Hasil saduran tersebut diterjemahkan lagi ke dalam bahasa asli (contoh bahasa Inggris) oleh orang yang memiliki kemampuan bahasa dan ahli dalam bidang aspek yang diukur; 4) Melakukan uji coba untuk memperoleh tingkat validitas dan reliabilitas dari instrumen yang disadur tersebut; 5) membuat norma; dan 6) Menyusun manual (Kartadinata dalam Helm a, 2001: 52).

Berkaitan dengan atribut psikologis individu dalam pengukuran, Suryabrata (1999: 67; 252) membagi ranah pengembangan alat ukur psikologis menjadi dua bagian yaitu pengembangan alat ukur atribut kognitif dan atribut non-kognitif. Pembagian itu dilakukan dengan asumsi bahwa atribut psikologis yang biasa dipersoalkan dalam psikologi tidak mempunyai eksistensi riil, dan hanya rekaan teoretis (*theoretical construct).* Oleh karena itu atribut psikologis tidak dapat diukur secara langsung, melainkan tidak langsung. Pengukuran tidak langsung itu maksudnya pengukuran yang dilakukan melalui respons yang dibuat oleh subjek pada waktu subjek dihadapkan kepada perangsang tertentu. Respons yang diperlukan untuk pengukuran atribut kognitif tidak sama dengan respons yang diperlukan untuk pengukuran atribut non-kognitif. Untuk pengukuran atribut kognitif diperlukan respons jenis pendapat *(judgement*), yaitu jenis respons yang dapat memunculkan benar atau salah. Sedangkan pengukuran atribut non-kognitif diperlukan respons jenis ekspresi sentimen (*expression of sentiment*), yaitu jenis respons yang tak dapat dinyatakan benar atau salah, atau seringkali dikatakan semua respons benar menurut alasannya masing-masing.

Berdasarkan pembagian atribut individu di atas berpengaruh pada spesifikasi tes yang dikembangkan pada penelitian ini. Suryabrata (1999: 86)

mengatakan bahwa kegiatan pengembangan spesifikasi alat ukur pada dasarnya adalah proses pengambilan keputusan. Setiap keputusan itu harus diambil berdasarkan pertimbangan-pertimbangan mengenai berbagai hal. Pertimbangan-pertimbangan tersebut pada umumnya mengenai hal-hal berikut seperti: a) Menentukan wilayah yang akan dikenai pengukuran; b) Menentukan dasar konseptual atau dasar teoretis yang akan digunakan sebagai landasan; c) Menentukan subjek yang akan dikenai pengukuran; d) Menentukan tujuan pengukuran; e) Menentukan materi alat ukur; f) Menentukan tipe soal yang akan digunakan; g) Menentukan tipe-tipe soal yang akan digunakan; h) Menentukan jumlah soal untuk keseluruhan alat ukur dan masing-masing bagiannya; i) Menentukan taraf kesukaran dan distribusi soal; j) Menyusun kisi-kisi atau "*test blue print*"; k) Melakukan perakitan soal; dan 1) Merencanakan jadwal penerbitan alat ukur.

Langkah-langkah pengembangan spesifikasi alat ukur atribut kognitif dan non-kognitif pada dasarnya sama, yaitu: a) Pengembangan spesifikasi alat ukur; b) Penulisan pernyataan atau pertanyaan; c) Penelaahan pernyataan atau pertanyaan; d) Perakitan instrumen (untuk keperluan uji coba); e) Uji coba; f) Analisis hasil uji coba; g) Seleksi dan perakitan instrumen; h) Pencetakan alat ukur; i) Administrasi instrumen (bentuk akhir); dan j) Penyusunan skala dan norma.

3. Pembakuan Alat Ukur

Menurut Suryabrata (1999: 5) pembakuan atau standirasasi diperlukan agar ilmuwan atau peneliti yang berbeda dan bekerja secara terpisah menghasilkan hasil yang sama atau sekurang-kurangnya setara. Anastasi & Urbina (2003: 4) mengatakan bahwa standarisasi mengimplikasikan keseragaman cara dalam penyelenggaraan dan penskoran tes. Senada dengan dua pendapat tersebut Gronlund (1985: 264) mengatakan *"...Standard content and procedure make it possible to give an identical test to individuals in different places at different times"*.

Dalam rangka menjamin keseragaman tersebut ada dua langkah penting yang harus dilakukan oleh pengembang. *Pertama*, menyediakan petunjuk-petunjuk yang rinci bagi penyelenggaraan alat ukur yang dikembangkan. Petunjuk-petunjuk itu meliputi jumlah materi yang digunakan, batas waktu yang disediakan untuk mengerjakan tes, instruksi-instruksi lisan, demonstrasi awal, cara menjawab pertanyaan dari peserta tes, dan setiap rincian lain dari situasi testing serta petunjuk-petunjuk lain yang ditujukan bagi penguji *{tester). Kedua,* adalah penetapan norma-norma yang mengindikasikan bagi siapa dan untuk tujuan apa tes tersebut cocok dipakai, sekaligus cara pelaksanaan dan penginterpretasiannya (Gronlund, 1985: 264; Anastasi & Urbina, 2003: 5). Bagaimanapun, kedua langkah standarisasi ini dimaksudkan untuk menyediakan petunjuk "baku" bagi tes yang bersangkutan, yang harus diikuti sepenuhnya oleh setiap pengguna tes tersebut (Subino dalam Helma, 2001: 53).

Dalam proses standarisasi, alat ukur yang dikembangkan perlu diselengarakan pada sampel yang luas dan representatif dari jenis orang yang memang menjadi sasaran perancangan alat ukur tersebut. Kelompok ini dikenal sebagai sampel standarisasi (*standardized sample*), yang berfungsi untuk menetapkan norma (Schmidt, 1999: 207; Anastasi & Urbina, 2003: 5). Nunnally (1978: 279) mengatakan *"...subjects should be used to obtain data for item analysis - five subjects per item should be considered the minimum that can be tolerated"*. Menurutnya banyak anggota sampel minimal (yang dapat ditolelir) bagi uji coba alat ukur dalam analisis butir soal adalah 5 kali lipat jumlah butir pernyataan yang akan dianalisis.

Di samping itu, terdapat dua tujuan dasar lagi yang harus diperoleh melalui standarisasi, yakni tingkat kesahihan (validitas) dan reliabilitas dari sebuah tes. Sebagaimana dikemukakan oleh Freeman I Helma, 2001: 53) bahwa *"...the fundamental purpose of standardizing a psychological test is to establish its reliability and its validity at as-high a level as possible"*.

a. Validitas

Menurut Suryabrata (1999: 56-57) validitas atau kesahihan digunakan sekurang-kurangnya dalam tiga konteks, yaitu: 1) validitas penelitian (*research validity*); 2) validitas soal *(item validity);* dan 3) validitas alat ukur atau tes (*test validity).* Pada validitas penelitian mempersoalkan derajat kesesuaian antara data hasil penelitian dengan keadaan sebenarnya; sejauhmana hasil penelitian mencerminkan keadaan yang sebenarnya Validitas ini mengandung dua sisi, yaitu validitas internal dan validitas eksternal. Untuk mendapatkan validitas internal penelitian, peneliti menggarapnya melalui penggunaan instrumen pengambil data yang memenuhi persyaratan tertentu. Validitas eksternal penelitian mempersoalkan derajat kesesuaian antara generalisasi hasil penelitian dengan keadaan yang sebenarnya.

Pada validitas soal perlu diingat bahwa validitas tes (alat ukur) bukan validitas soal. Validitas soal adalah derajat kesesuaian antara suatu soal dengan perangkat soal-soal yang lainnya. Ukuran validitas soal adalah korelasi antara skor pada soal itu dengan skor pada perangkat soal (*item-total correlation).* Isi validitas soal adalah daya pembeda soal *(item discrimination power*) bukan validitas tes. Jadi, seorang pengembang alat ukur tidak dapat mengklaim telah memiliki validitas tes karena memiliki kumpulan validitas soal.

1) Validitas alat ukur (tes)

Validitas alat ukur menyangkut apa yang diukur suatu alat ukur dan seberapa baik alat ukur itu bisa mengukur hal yang akan diukur (Anastasi & Urbina, 2003: 85). Validitas ini ditentukan melalui pendapat profesional *(professional judgment*) dalam proses telaah soal (item). Dengan menggunakan spesifikasi alat ukur yang dikembangkan (telah ada) pengembang melakukan analisis logis untuk menetapkan apakah butir-butir soal *(item)* yang telah dikembangkan memang mengukur (representatif bagi) apa yang dimaksud untuk diukur.

2) Validitas Konstruksi Teoretis (*Construct Validity)*

Validitas konstruksi teoretis mempersoalkan sejauh mana skor-skor hasil pengukuran dengan instrumen yang dipersoalkan itu merefleksikan konstruksi

teoretis yang mendasari penyusunan alat ukur tersebut. Validitas berdasarkan konstruksi teoretis ini merupakan proses kompleks yang memerlukan analisis logis dan dukungan data empiris. Menurut Suryabrata (1999: 60-66) ada dua metode yang diakui oleh para ahli untuk mendapatkan validitas konstruk. a) Analisis faktor

Analisis faktor merupakan analisis statistik yang bertujuan untuk mengidentifikasi, mengelompokkan, dan meringkas faktor-faktor yang merupakan dimensi suatu variabel, definisi dan sebuah fenomena tertentu (Nugroho, 2005: 91). Pada prinsipnya analisis faktor digunakan untuk mereduksi data, yaitu untuk meringkas sejumlah variabel menjadi lebih sedikit dan menamakannya faktor.

Di dalam analisis sebuah alat ukur, analisis faktor diterapkan untuk menyederhanakan butir-butir soal menjadi satu atau beberapa faktor yang merupakan inti yang melandasi konstruk alat ukur tersebut. Faktor-faktor yang dikandung oleh butir soal lazimnya diturunkan dari korelasi antar item.

Suryabrata, 1999: 58). Menurut Arikunto (2002: 144-145) suatu instrumen atau alat ukur dikatakan valid apabila mampu mengukur apa yang diinginkan; mengungkap data dari variabel yang diteliti secara tepat. Tinggi rendahnya validitas alat ukur menunjukkan sejauhmana data yang terkumpul tidak menyimpang dari gambaran tentang validitas yang dimaksud. Intinya, validitas alat ukur mencerminkan ketepatan suatu alat ukur untuk mengukur apa yang ingin diukur.

Suryabrata (1999: 58) mengatakan bahwa secara konvensional validitas alat ukur dapat dilihat dari tiga arah, yaitu: 1) Dari isi yang hendak diukur; 2) Dari arah rekaan teoretis (*construct*) atribut yang diukur, dan 3) Dari kriteria alat ukur. Oleh karena itu validitas alat ukur biasanya dibedakan menjadi tiga macam, sebagaimana terlihat pada gambar di bawah ini.

b. Reliabilitas

Reliabilitas alat ukur menunjukkan sejauh mana derajat keajegan (konsistensi) skor yang dicapai oleh testi dari suatu pengukuran dengan alat ukur

yang sama pada kondisi yang berbeda. Dengan kata lain, reliabilitas alat ukur merujuk pada sejauh mana perbedaan-perbedaan skor perolehan mencerminkan perbedaan-perbedaan atribut sebenarnya

Kaitannya dengan reliabilitas skor, duduk perkaranya terletak pada definisi varian kesalahan. Kondisi apapun yang tidak relevan dengan maksud alat ukur merupakan varian kesalahan. Anastasi & Urbina (2003: 63) mengatakan bahwa bagaimanapun kondisi pengukuran diusahakan maksimum, tak ada satu pun alat ukur yang dapat dipercaya. Hal ini menuntun kepada definisi dasar reliabilitas sebagai proporsi varians skor perolehan yang didapat dari penjumlahan antara varians skor murni dengan varians kekeliruan pengukuran.

Karena reliabilitas alat ukur itu berkenaan dengan derajat konsistensi atau kesamaan antara dua perangkat skor, maka semua jenis reliabilitas dinyatakan dalam bentuk koefisien korelasi (r) (Suryabrata, 1999: 41; Anastasi & Urbina, 2003: 64). Besar kecilnya reliabilitas suatu alat ukur ditentukan oleh besar kecilnya nilai korelasi hasil tes yang dinamakan *indeks reliabilitas.*

Perbedaan pengertian reliabilitas sangat bergantung pada bagaimana indeks reliabilitas dihitung. Paling tidak terdapat empat konsep reliabilitas, yaitu: 1) *Paralel* atau *ekuivalen*; 2) *Test-retest* atau *stabilitas*; 3) *Split-half* atau *belah dua;* dan 4) *Internal consistency.* Sebagaian ahli berpendapat bahwa metode *split-half* atau *belah dua* merupakan bagian dari metode keajegan internal (*internal consistency)* sehingga pembagian metode menjadi tiga bagian, yaitu: 1) *Test-retest;* 2) *Paralel:*; dan 3) *Internal consistency.* Berhubungan dengan itu, Suryabrata (1999: 42) mengatakan bahwa reliabilitas alat ukur yang juga menunjukkan derajat kekeliruan pengukuran tak dapat ditentukan dengan pasti, melainkan hanya dapat diestimasi. Menurutnya ada tiga pendekatan yang biasa digunakan dalam mengestimasi reliabilitas alat ukur, yaitu: 1) Pendekatan tes ulang; 2) Pendekatan tes paralel; dan 3) Pendekatan satu kali pengukuran.

c. Norma

Berdasarkan tulisan Anastasi & Urbina (2003: 36-49), dikatakan bahwa skor-skor pada pengukuran psikologis paling umum diinterpretasikan dengan

mengacu pada *norma-norma* yang menggambarkan kinerja tes dari sampel standarisasi. Dengan demikian norma-norma secara empiris ditetapkan dengan menentukan apa yang sesungguhnya dilakukan oleh orang-orang dalam kelompok yang representatif. Skor mentah individu selanjutnya dibandingkan dengan distribusi skor yang diperoleh melalui sampel standarisasi untuk menemukan di mana letaknya dalam distribusi itu.

Dalam rangka menilai posisi individu berdasarkan pada sampel standarisasi skor mentah diubah menjadi ukuran relatif. Skor yang diturunkan tersebut dirancang untuk melayani dua fungsi, yaitu: 1) Menunjukkan sikap positif relatif individu dalam sampel normatif, sehingga dengan begitu dimungkinkan satu evaluasi tentang kinerjanya dalam acuan individu lain; dan 2) Memberikan ukuran yang dapat dibandingkan sehingga dimungkinkan untuk dilakukan perbandingan langsung dari kineija individu pada berbagai tes.

Ada berbagai cara mengkonversikan skor mentah untuk memenuhi dua fungsi di atas. Pada dasarnya, skor-skor yang dihasilkan diungkapkan dengan dua cara pokok, yaitu: 1) Tingkat perkembangan yang dicapai; dan 2) Posisi relatif dalam kelompok khusus. Cara pertama dikenal dengan norma-norma perkembangan, yaitu memberi arti pada skor-skor tes dengan mengindikasikan sejauhmana seorang individu telah maju sepanjang jalur perkembangan normal. Biasanya norma perkembangan ini terbagi berdasarkan usia mental, ekuivalen kelas *(grade equivalents*), dan skala ordinal. Cara kedua dikenal dengan norma-norma dalam kelompok dimana kinerja individu dievaluasi dari segi kinerja kelompok standarisasi yang paling bisa dibandingkan seperti membandingkan skor mentah anak-anak yang berusia kronologis sama atau berada dalam kelas di sekolah yang sama.

Skor kelompok yang dimaksud sebelumnya memiliki arti kuantitatif vang seragam dan dirumuskan secara jelas serta dapat digunakan dengan cukup sesuai dengan kebanyakan jenis analisis statistik. Norma-norma dalam kelompok ini adalah persentil dan skor standar seperti skor z dan skor t (Anastasi & Urbina. 2003: 44-49).

4. Beberapa Teknik Analisi Kebaikan Tes

a. Analisis Validitas Tes

Secara psikologis, bermakna-tidaknya skor-skor sebuah tes tergantung kepada berkorelasinya dengan skor-skor tes lainnya. Dapat saja terjadi bahwa sebuah tes yang memiliki koefisien keterandalan positif sempurna (r_{tt}= 1,00) tidak berkorelasi sama sekali dengan skor-skor tes lainnya. Tes semacam itu tidak memiliki nilai praktis sama sekali oleh karena skor-skornya tidak mempunyai hubungan lain kecuali dengan dirinya sendiri. Dengan perkataan lain tes seperti itu tidak dapat ditafsirkan sama sekali oleh karena tidak dapat ditera dengan menggunakan tolok ukur yang lain. Tes semacam itu dapat dikatakan mantap akan tetapi tidak tepat.

Validitas tes dapat didefinisikan sebagai proporsi varians yang sebenarnya yang mempunyai hubungan dengan tujuan tes. Mempunyai hubungan di sini diartikan sebagai berasal dari faktor atau faktor-faktor yang diukur oleh tes yang bersangkutan. Faktor di sini dapat berupa inteligensi, bakat khusus, kepribadian, atau ka**r**akteristik-karakteristik lainnya. Oleh sebab itu reliabilitas tes itu dapat didefinisikan dengan salah satu dari definisi berikut. Pertama, sebagai derajat kebaikan tes yang menjelaskan seberapa jauh tes tersebut mampu mengukur faktor yang melandasi karakteristik-karakteristik yang dihipotesiskan**,** dan yang kedua sebagai hubungan dengan skor-skor tes yang lain yang bertindak sebagai tol**a**k ukurnya.

Di dalam pembakuan tes, pembaku tes dapat memilih salah satu dari empat macam keshahihan tes yang cocok bagi tes yang sedang dibakukan. Macam validitas tersebut adalah validita**s** bangun (*construct validity*), validitas isi (*content validity*), validitas tolok ukur (*criterion validity*), dan validitas ramal (*predictive validity*) (Thorndike and Hagen, Jakson 2003). Validitas jenis pertama validitas bangun (*construct validity*), menyatakan *traits* apa yang diukur oleh **t**es yang ber saungkutan. Oleh sebab itu menganalisis validitas ini berarti menganalisis apakah tes yang bersangkutan mengukur apa yang seharusnya diukur. Sasaran analisis

diarahkan kepada *traits* yang sedang diukur oleh tes yang bersangkutan. Oleh sebab itu analisisnya dilakukan secara empiris sehingga hasilnya berupa sebuah index kuantitatif.

validitas isi (*content validity*), menyatakan apakah butir-butir soal sebuah tes mewakili apa yang hendak diukur oleh tes yang bersangkutan. Sasaran analisis keshahihan macam kedua ini diarahkan kepada tugas atau tingkah laku yang dituntut kepada testi didalam menyelesaikan tes yang bersangkutan. Tidak seperti pada keshahihan macam pertama maka analisis keshahihan macam ke dua ini dilakukan secara rasional. Oleh sebab itu hasilnya juga tidak berupa suatu index kuantitatif.

Validitas tolok ukur (*criterion validity*), dan validitas ramal (*predictive validity*) menyatakan seberapa tepat prestasi yang berhasil dicapai di dalam sebuah tes dapat meramalkan prestasi di dalam tes yang lain yang berdiri sendiri. Analisis di dalam keshahihan macam ketiga dan keempat ini dilakukan dengan jalan membandingkan antara skor-skor tes yang bersangkutan dengan skor-skor tes yang diramalkan. Sasaran ditujukan kepada faktor atau variabel yang diramalkan yang bertindak sebagai tolok ukurnya. Keshahihan tes macam ke tiga dan keempat ini dianalisis secara korelasional sehingga hasilnya berupa suatu index kuantitatif.

Berdasarkan uraian singkat di atas maka dapat dinyatakan bahwa bagi tes psikologis harus memiliki validitas bangun yang mantap dan validitas ramal yang tinggi. Alasannya adalah oleh karena setiap tes psikologis memiliki tugas yang jelas tentang apa yang mampu diukurnya.

Koefisien Validitas tes dapat ditafsirkan ke dalam berbagai makna. Brown (Jakson 3003) menyebutkan dua makna, yakni sebagai derajat ketepatan tes dan sebagai proporsi varians. Sedangkan Brogden (Anastasi & Urbina 2003:8) memberi makna kepada koefisien keshahihan tes itu sebagai keampuhan meramal. Ia memperlihatkan, sebagai dukungan atas pendapatnya itu, bahwa koefisien keshahihan sebuah tes itu sama dengan rasio skor rata-rata n peserta tes yang dipilih dengan tes tersebut dengan skor rata-rata n peserta tes yang akan dipilih dengan tes yang sama. Namun demikian, koefisien

keshahihan tes yang dimaknakan sebagai keampuhan meramal, yang didekati dengan teknik korelasional itu belum tuntas oleh karena belum dapat membuahkan kekeliruan ramal. Ia belum dapat menjawab persoalan bagaimana skor-skor pada variabel kriteria itu tergantung kepada skor-skor pada variabel prediktor **y**ang dapat menuntaskan analisis ini adalah persamaan regresi, yang jika bersifat linier dapat dinyatakan dengan $\hat{Y}$ = a + bx. Dengan persamaan tersebut maka Ŷ setiap testi dapat dihitung**,** begitu juga dengan kekeliruannya, yakni selisih antara Y dengan $\hat{Y}$. Hubungan antara koefisien keshahihan dengan regresi ini dapat dinyatakan sebagi berikut : makin Tinggi koefisien validitas, makin berdekatanlah kedudukan antara skor-skor pada variabel kriteria dengan skor- skor p**a**da variabel prediktor. Atau dengan kata lain dapat pula dinyatakan bahwa makin shahih sebuah **t**es maka makin tepat daya ramalnya. Oleh sebab itu maka korelasi antara Y dengan $\hat{Y}$ itu sama dengan hubungan antara X dengan Y (Huck, Cormier, and Bounds, Subino 1984**).**

b. Analisis Reliabilitas Tes

Secara harfiah maka tingkat reliabilitas tes dapat dinyatakanan sebagai tingkat ketetapan skor-skor yang diperoleh testi dari suatu pengetesan. Konsep reliabilitas tes seperti yang dinyatakan di atas bertumpu pada skor yang sebenarnya, *the true score*, yakni skor mentah yang dicapai oleh testi, sehingga skor tersebut mampu mengukur karakteristik-karakteristik testi tanpa kekeliruan. Akan tetapi pada kenyataannya hal itu tidak lah semudah apa yang kita bayangkan, karena pada kenyataanya tidak ada tes yang mampu mengukur karakteristik testi tanpa kekeliruan. Oleh sebab itu skor yang dicapai seseorang di dalam suatu tes sebetulnya terdiri atas skor yang sebenarnya ditambah dengan skor yang diperoleh karena kekeliruan yang terjadi di dalam pengetesan, yang secara operasional dapat dinyatakan dengan X_t = + E (Brown, Subino 1984:115). Atau dapat pula dinyatakan di dalam bentuk varians menjadi $s^2_t = s^2_T + s^2E$ dengan asumsi bahwa T dan E tidak saling bergantungan (*independent*) Model ini dapat dipetakan ke dalam sebuah grafik sebagai berikut.

Satu masalah penting yang perlu dibahas lebih lanjut ialah masalah sumber kekeliruan di dalam pengetesan. Perlunya hal tersebut dibahas lebih lanjut adalah karena hal tersebut menjadi dasar prosedur dan teknik menaksir keterandalan tes. Brown (Jakson 2003) menyebutkan bahwa yang menjadi sumber-sumber kekeliruan tersebut adalah tes yang bersangkutan, pelaksanaan tes, dan peserta tes. Tes seoagai sumber kekeliruan dapat terjadi karena kekaburan kalimat butir-butir soalnya, karena kekaburan petunjuk pengerjaannya, dan karena kekeliruan penyekorannya. Pelaksanaan tes sebagai sumber kekeliruan dapat terjadi karena salah memberikan petunjuk pengerjaan tes, salah memberi waktu (terutama pada tes kecepatan), dan tidak teliti mem-perhatikan hal-hal yang dapat mengganggu jalannya tes.

Peserta tes juga dapat menjadi sumber kekeliruan, contohnya tes akan mengalami bias dalam hasilnya apabila motivasi testi yang rendah, karena pengaruh pengalaman dengan tes yang bersangkutan, karena kecemasan yang menghinggapi testi selama tes berlangsung, dan karena kesehatan testi. Di dalam penelitian ini, yang diduga dapat menjadi sumber kekeliruan tersebut terbatas karena kekurang-jelasan di dalam memberi petunjuk cara mengerjakan tes, motivasi testi yang rendah, dan karena kecemasan. Oleh sebab itu skor-skor tes yang di diperoleh para testi diduga dapat diandalkan. Namun demikian semuanya itu masih akan dibuktikan oleh hasil uji coba yang dilakukan di dalam penelitian ini.

Untuk mengurangi kemungkinan terjadinya kekeliruan-kekeliruan seperti yang diutarakan di atas, maka idealnya analisis keterandalan tes itu harus dilakukan berkali-kali. Namun yang ideal itu sangat sukar untuk dilaksanakan. Sebagai gantinya maka analisis keterandalan tes tersebut cukup didasarkan atas satu atau dua kali pengetesan saja.

Atas dasar teknik yang digunakan, maka terkembangkanlah tiga macam keterandalan tes, yakni koefisien kemantapan (*coefficient of stability*), koefisien ekuivalens (*coefficient of eguivalence*), dan koefisien ketetapan dalam (*coefficient of internal conslstency*). Yang pertama diperoleh dengan jalan

mengorelasikan dua perangkat skor hasil dua kali pengetesan dengan selang waktu tertentu dan menggunakan tes serta testi uji coba yang sama. Yang kedua diperoleh dengan jalan mengorelasikan dua perangkat skor hasil dua tes sejajar yang diteskan kepada testi uji coba yang sama dengan selang waktu tertentu. Dan yang ketiga diperoleh dengan jalan mengorelasikan dua perangkat skor dari dua parohan tes, atau dengan teknik Kuder-Richardson 20 (KR20).

Jika diperhatikan dengan saksama, maka koefisien kemantapan dan koefisien ekuivalen tersebut baru menunjukkan derajat ketetapan skor-skor tes. Koefisien-koefisien tersebut sama sekali belum mencerminkan kemungkinan banyaknya faktor yang dimuat oleh tes yang bersangkutan. Koefisien keterandalan yang sudah mencerminkan banyaknya faktor yang dimuat oleh tes yang bersangkutan adalah koefisien ketetapan dalam. Dasar pemikirannya adalah, walaupun koefisien ketetapan dalam tidak identik dengan proporsi muatan faktor yang dikandung oleh sesuatu tes, akan tetapi koefisien ketetapan dalam yang tinggi merupakan syarat mutlak bagi tes yang bersangkutan agar ia hanya mengukur satu faktor saja.

Makna keterandalan tes seperti yang telah diuraikan di atas tidak lebih dari seberapa jauh skor-skor sebuah tes itu dapat dipercaya. Sedangkan masalah seberapa besar kekeliruan skor yang dicapai oleh seseorang testi itu dapat ditaksir, sama sekali belum dapat diketahui. Padahal masalah kekeliruan skor atau besarnya kemungkinan berubahnya (*fluctuation*) skor yang dicapai seseorang di dalam suatu tes dari waktu ke waktu itu merupakan masalah yang sangat penting bagi siapa pun yang berkepentingan dengan skor tes.

c. Teknik Menganalisis Daya Pembeda

Koefisien Validitas dan Reliabilitas tes adalah derajat kebaikan tes secara keseluruhan. Kedua ukuran statistik tersebut sama sekali belum menjelaskan derajat kebaikan atau butir-butir soal yang membangun tes tersebut dapat diketahui antar lain melalui analisis daya pembedanya.

Ada dua teknik uji daya pembeda butir-butir soal tes. Pertama adalah dengan jalan menganalisis perbedaan keberhasilan menjawab dengan betul setiap butir soal antara kelompok unggul dengan kelompok asor. Jalan fikiran yang melandasi teknik ini ialah, jika kedudukan skor total seseorang testi berada pada kelompok unggul, maka kedudukan skor setiap butir soalnya juga berada pada kelompok unggul. Begitu pula dengan kelompok asor. Jadi secara teoretis maka kelompok unggul dengan kelompok asor tersebut berbeda secara signifikan. Oleh sebab itu butir-butir soal yang dapat membedakan secara signifikan keberhasilan menjawab antara kelompok unggul dengan kelompok asor dapat dinyatakan memiliki daya pembeda yang signifikan.

Teknik uji daya pembeda yang ke dua ialah dengan jalan mengkorelasikan antara skor setiap butir soal dengan skor total. Jalan fikiran yang melandasi teknik ini ialah, bahwa tes secara keseluruhan dengan masing-masing butir soal yang membangun tes tersebut mengukur kawasan atau karakteristik yang sama. Oleh sebab itu secara teoretis keduanya juga berkorelasi secara signifikan.

Teknik analisis yang pertama tersebut telah dikembangkan oleh Guilforad dengan grafik abac-nya; Izard (Nugroho, 2005: 48) dengan mengadaptasikan grafik abac Turner dan melengkapi nya dengan tabel tolak ukur signifikansi dikaitkan dengan banyaknya testi; dan Ross bersama-sama dengan Stanley (Nugroho, 2005: 92) yang mengembangkan tabel batas kritis signifikansi daya pembeda setiap butir soal tes menurut bentuknya.

d. Teknik Mengenalisis Tingkat Kesukaran Butir soal Tes

Untuk menganalisis tingkat kesukaran butir-butir soal tes bentuk pilihan berganda dapat digunakan tenik-teknik klasik, unggul - asor, kesukaran relatif, dan rasio kesukaran relatif.

Teknik klasik yang menghitung tingkat kesukaran butir- butir soal bentuk pilihan berganda dengan jalan membagi banyaknya testi yang dapat menjawab dengan betul butir soal yang ke-1.

Teknik menggunakan kelompok unggul - asor, seperti yang pernah dikembangkan oleh Izard (Nugroho, 2005: 32 - 33) mengembangkat tingkat

kesukaran dengan cara membedaka skor- skor yang didapatkan oleh peserta tes yang dikategorikan terlebih dahulu pada kategori unggul dan asor untuk kemudian dijadikan untuk dasar perhitungan tingkat kesulitan pada butir soal tertentu. Pada prinsipnya teknik ini menganalisis skor suatu soal berdasarkan skor yang didapatkan oleh kelompok unggul dan kelompok asor pada suatu pelaksanaan tes, adapun notasi dari rumus tersebut adalah sebagai berikut:

$$TK = \frac{BA+BB}{NA+NB} \textbf{ X 100\%}$$

TK = indeks tingkat kesukaran satu butir soal trtentu

BA = Jumlah siswa yang menjawab benar pada kelompok A

BB = Jumlah siswa yang menjawab benar pada kelompok

NA = Jumlah siswa pada kelompok A

NB = Jumlah siswa pada kelompok B

Teknik rasio kesukaran realtif ini dikembangkan oleh David A. Frisbie (Nugroho, 2005: 75**)** dengan maksud untuk menjadikan rasio Kesukaran relatif itu sebagai index kesukaran butir-butir soal sebuah tes dan index kesukaran tes sacara keseluruhan. Keunggulan teknik ini ialah, pertama tidak menuntut terpenuhinya asumsi bahwa distribusi skor tes yang bersangkutan normal. Kedua, teknik ini memungkinkan dihitung nya tingkat kesukaran butir-butir soal sebuah tes yang terdiri atas bagian-bagian tes yang memiliki alternatif jawaban yang tidak sama banyaknya, yang tentunya bagi sebuah tes yang terbangun dari butir-butir soal yang memiliki alternatif jawaban yang sama.

5. Kecenderungan Baru dalam Pembakuan Tes

Kecenderungan baru di dalam membakukan tes ini terjadi di dua sisi. Sisi pertama adalah landasan teoretis» dan sisi ke dua adalah teknik analisis. Dari sisi landasan Teoritis ada kecenderungan untuk menganalisis tes tidak semata-mata berdasarkan skor-skor hasil uji coba tes yang bersangkutan akan tetapi juga dikaitkan dengan faktor-faktor latar belakang kehidupan testi uji coba. Ini di-maksudkan agar penggunaan tes yang sedang dibakukan tersebut bagi kalangan yang lebih luas lagi tidak bias oleh faktor- faktor yang melatarbelakangi kehidupan testi. Contoh dari kecenderungan ini misalnya analisis terhadap sebuah

tes inteligensi yang ditelaah juga dari besarnya urunan faktor-faktor bukan intelektual terhadap skor tes sehingga penafsirannya menjadi lebih tuntas. Tuntasnya penafsiran tersebut disebabkan karena skor tes selalu dikaitkan dengan faktor-faktor lain yang melatarbelakangi kehidupan testi sehingga testi diperlakukan sebadai sesuatu yang utuh.

Dari sisi teknik analisis, kecenderungan tersebut ialah timbulnya pandangan bahwa melibatkan analisis faktor ke dalam analisis tes merupakan suatu keniscayaan. Dengan makin jelasnya kelemanan- kelemahan yang terlihat pada koefisien keterandalan dan Koefisien kesnahihan sebadai ukuran-ukuran yang menyatakan kebaikan tes, maka kecenderungan untuk melibatkan analisis faktor ke dalam analisis tes makin kuat. Alasannya sangat sederhana akan tetapi sangat mendasar, yakni bahwa dengan analisis faktor maka dapat diketahui hakekat kemampuan ukur tes. Di samping itu alasan-alasan kemajuan teknologi juga mendukungnya.

D. Multidimensional Aptitude Battery (MAB - II)

1. Deskripsi Umum Tes MAB II

Multidimensional Aptitude Battery-II (MAB-II) merupakan tes yang dirancang untuk memberikan pengukuran yang lebih objektif dan terandal terhadap kemampuan kognitif umum atau dikenal dengan inteligensi, dalam bentuk profil yang mencakup skor dari lima sub-tes verbal dan lima sub-tes performa.

Tes MAB-II dirancang untuk kepentingan asesmen kemampuan intelektual secara luas (mencakup banyak responden), bagi orang dewasa maupun remaja usia 16 tahun ke atas. Sama halnya seperti pengukuran kemampuan intelektual lainnya, tes ini dapat digunakan untuk beragam tujuan dalam beragam konteks. Contohnya, dalam lingkup konseling karir dan kependidikan, industri dan bisnis, fasilitas klinis dan kesehatan mental, serta untuk penelitian tingkat

dasar. Untuk menghasilkan fungsi assessment yang memadai, pelaksanaan tes MAB-II memerlukan keterampilan bahasa yang baik untuk membaca maupun memahami pentunjuk tertulis serta untuk menguasai petunjuk lisan di dalamnya. Tidak dianjurkan untuk digunakan terhadap responden dengan retardasi mental, atau individu yang memiliki gangguan psikotik, karena berpengaruh terhadap pemahaman dan pelaksanaan intruksi tes. Meskipun demikian, berdasarkan pengalaman penulis, banyak pasien penderita gangguan psikotik (yang tidak terlalu parah) dapat menyelesaikan tes MAB-II.

Dalam aturan penyekalaan tes MAB-II memiliki dua kelompok skala yakni verbal dan performa. Kelima sub-tes verbal tersaji dalam satu buku tes dan kelima sub-tes performa tersaji dalam buku tes lainnya. Lembar jawaban tes terpisah (baik untuk penyekoran manual maupun dengan mesin) diberikan pada masing-masing pelaksanaan sub-tes, sehingga memungkinkan untuk mengadministrasikan sub-tes verbal dan sub-tes performa secara terpisah ataupun secara bersamaan. Skor dari kedua sub-tes dan skor dari kesepuluh sub-skala dalam MAB-II dapat menghasilkan IQ verbal, IQ performa, dan IQ total (full-scale IQ). Skor standar dapat digunakan untuk mewakili IQ verbal, performa maupun total.

Seluruh sub-skala dalam MAB-II secara substansial berkaitan dengan kemampuan intelektual secara umum, yang secara konsisten penting untuk memprediksi kinerja dalam semua ranah pekerjaan. Pola tinggi rendahnya skor tes dapat membantu untuk konseling atau penempatan individu dalam suatu pekerjaan yang memungkinkan mereka untuk menyadari potensi mereka di bidang tertentu.

2. Sejarah dan Perkembangan Tes MAB II

Sejarah tes MAB-II diawali oleh akar peristiwa yang sama dengan tes inteligensi yang pertama kali di publikasikan oleh Binet dan Simon pada tahun 1905.Penelitian Binet dan Simon menginspirasi standarisasi formal dari tes inteligensi yang sangat populer yakni Stanford-Binet oleh Terman (1916).

Pengembangan tes MAB-II juga diilhami oleh perkembangan pengukuran inteligensi non-verbal terhadap individu, seperti pada studi Pintner dan Anderson (1917) dan Healy (1921) yang menghasilkan tes melengkapi gambar, serta studi Kohs (1923) yang menghasilkan tes merancang balok susun. Pintner dan Paterson (1917) mengembangkan dan menyetandarkan skala performansi multitask individu yang pertama.

Implementasi tes kemampuan kognitif kelompok modern pertama kali dilakukan oleh psikolog yang mengembangkan Tes Klasifikasi Personil Angkatan Darat Amerika Serikat. Implementasi tes tersebut berlandaskan pada dua studi penting yakni Skala Performa yang disusun oleh Cornell dan Coxe (1934) serta Skala Inteligensi yang disusun oleh Wechsler dan Bellevue (1939).

Pengukuran multi-skala terhadap kemampuan intelektual dalam tradisi Alfa dan Beta Angkatan Darat Amerika Serikat telah berhasil dikarenakan oleh beberapa alasan, yakni : (a) Tes tersebut menggabungkan tugas-tugas yang beragam, tidak hanya konten verbal dan pengetahuan umum tapi juga performan dan keterampilan praktis; (b) Konten tes sesuai bagi anak-anak, remaja dan orang dewasa; (c) Skor tes merefleksikan konsepsi terbaru mengenai inteligensi dimana proses psikotik, kerusakan neurologis, atau gangguan emosional terlibat di dalamnya; (d) Tes memiliki penormaan yang teliti dan terstandarisasi; (e) Tes dapat menstimulasi munculnya penelitian-penelitian penting, artikel penelitian mengenai pengukuran inteligensi semakin banyak ditemukan; (f) Banyak tes inteligensi menunjukkan kualitas psikometrik yang tinggi; dan (g) Data validitas substansial untuk tes inleigensi banyak dihasilkan sepanjang tahun-tahun berikutnya, beragam mulai dari prediksi jenjang akademik sampai pada kualitas kinerja dan penyesuaian kerja (Anastasi dan Urbina, 1997).

Faktanya, skala inteligensi yang terstandarisasi telah menjadi kriteria bagi penilaian keberfungsian intelektual. Contohnya, apabila peringkat yang diberikan seorang guru terhadap siswanya bertolak belakang dengan hasil pengukuran inteligensi siswa tersebut, maka para psikolog akan mempertanyakan metode pemberian peringkat yang diberikan oleh guru tersebut.

Terdapat setidaknya satu kelemahan dalam mengaplikasikan tes inteligensi multi-skala pada lingkup yang lebih luas, karena tes tersebut memerlukan pengadministrasian secara individual dan penyekoran oleh professional yang terlatih secara khusus. Adalah benar bahwa pada situasi tertentu, observasi klinis terhadap perilaku responden saat mengerjakan tes akan sangat membantu, karenanya bagi responden yang mengalami tekanan mental atau gejala psikotik lainnya pelaksanaan tes haruslah individual. Namun, untuk memperkirakan keberfungsian intelektual responden dalam jumlah banyak, pelaksanaan secara individual sangat memakan biaya dan tidak terlalu diperlukan. Ketika pelaksanaan tes kemampuan intelektual yang dilakukan secara berkelompok dibandingkan dengan skala Wechsler (yang dilaksanakan secara individual) untuk memprediksi criteria terkait prestasi akademik, maka diperoleh data bahwa pelaksanaan tes secara individual pada umumnya tidak menunjukkan validitas predktif yang lebih baik dibandingkan dengan pelaksanaan secara berkelompok (Matarazzo, 1972). Hal ini membawa pada pertanyaan mengenai apakah mungkin untuk menggabungkan beberapa fitur positif dari tes multi-skala yang dilaksanakan secara individual ke dalam format tes terstruktur yang dilaksanakan secara berkelompok, dengan instruksi yang otomatis, dan dengan penyekoran manual maupun dengan mesin. Pertanyaan itulah yang kemudian membawa pada penyusunan tes MAB-II ini, untuk mengevaluasi sejauh mana kemungkinan penggabungan tersebut dapat dicapai.

Tes MAB-II ini memiliki tujuan yang sama seperti tes sebelumnya (MAB) yakni untuk memberikan pengukuran yang lebih objektif dan terandal terhadap inteligensi. Sebagian besar revisi pada edisi terbaru ini terletak pada pengenalan norma-norma baru, tabel norma yang mencakup nilai persentil, dan profil tes yang berdasarkan pada standarisasi yang sangat teliti. Selain dari ketiga hal itu fitur lain dalam MAB-II tetap tidak berubah, masih sama seperti edisi sebelumnya. Buku tes, lembar jawaban, kunci jawaban, petunjuk pengerjaan tes, dan aturan waktu tidak mengalami perubahan yang signifikan sejak edisi pertama.

Format laporan hasil tes MAB-II yang mencakup diagram konversi berdasarkan penormaan MAB-II berbeda dengan format laporan hasil tes sebelumnya. Pedoman tes yang baru juga memberikan deskripsi skala dan saran pekerjaan yang relevan dengan skor tertinggi pada masing-masing sub-skala.

Tujuan dilakukannya revisi terhadap MAB-II adalah : (a) Untuk menghasilkan pengukuran kemampuan kognitif umum yang berdasarkan atas norma Amerika Utara yang modern, merefleksikan karakteristik demografis yang sesuai zaman; (b) Untuk memisahkan ikatan antara MAB dan WAIS-Revised (Wechsler, 1981), kedua tes tersebut sebelumnya terhubung oleh kalibrasi statistik atau penyetaraan norma antara keduanya; dan (c) Untuk menggambarkan perkembangan dalam hal pelaporan skor, MAB-II memungkinkan penggunaan komputer untuk penyekoran dan pelaporan hasil tes.

3. Karakteristik Umum Tes MAB II

a. Kegunaan tes MAB II

Tes MAB-II dirancang untuk kepentingan assesmen kemampuan intelektual secara luas (mencakup banyak responden), bagi orang dewasa maupun remaja usia 16 tahun ke atas. Sama halnya seperti pengukuran kemampuan intelektual lainnya, tes ini dapat digunakan untuk beragam tujuan dalam beragam konteks. Contohnya, dalam lingkup konseling karir dan kependidikan, industri dan bisnis, fasilitas klinis dan kesehatan mental, serta untuk penelitian tingkat dasar. Untuk menghasilkan fungsi assessment yang memadai, pelaksanaan tes MAB-II memerlukan keterampilan bahasa yang baik untuk membaca maupun memahami pentunjuk tertulis serta untuk menguasai petunjuk lisan di dalamnya. Tidak dianjurkan untuk digunakan terhadap responden dengan retardasi mental, atau individu yang memiliki gangguan psikotik, karena berpengaruh terhadap pemahaman dan pelaksanaan intruksi tes. Meskipun demikian, berdasarkan pengalaman penulis, banyak pasien penderita gangguan psikotik (yang tidak terlalu parah) dapat menyelesaikan tes MAB-II.

Terkait translasi tes ke dalam bahasa selain bahasa Inggris, penulis menyediakan terjemahan resmi tes berbahasa Perancis dan Spanyol. Sub-tes verbal dalam MAB-II tidak disarankan untuk mengukur inteligensi individu yang bahasa pertamanya bukan bahasa Inggris, meskipun tetap dapat digunakan untuk mengevaluasi tingkat keberfungsian responden terkait bahasa yang digunakan dalam tes atau bahasa Inggris.

b. Aturan penyekalaan

Seperti halnya skala dari banyak pengukuran kemampuan inteligensi, tes MAB-II memiliki dua kelompok skala yakni verbal dan performa. Kelima sub-tes verbal tersaji dalam satu buku tes dan kelima sub-tes performa tersaji dalam buku tes lainnya. Lembar jawaban tes terpisah (baik untuk penyekoran manual maupun dengan mesin) diberikan pada masing-masing pelaksanaan sub-tes, sehingga memungkinkan untuk mengadministrasikan sub-tes verbal dan sub-tes performa secara terpisah ataupun secara bersamaan. Skor dari kedua sub-tes dan skor dari kesepuluh sub-skala dalam MAB-II dapat menghasilkan IQ verbal, IQ performa, dan IQ total (full-scale IQ). Skor standar dapat digunakan untuk mewakili IQ verbal, performa maupun total.

4. Asal Usul skala Verbal dan Performa pada Tes MAB

Merupakan suatu kebanggaan tersendiri untuk mengakui kontribusi David Wechsler di bidang pengukuran inteligensi orang dewasa dan khususnya sebagai rujukan utama bagi pengembangan tes MAB. Format tes dan nama skala yang digunakan oleh skala Wechsler baik verbal maupun performa, berasal dari tes Alfa-Beta yang diadministrasikan terhadap angkatan darat Amerika Serikat. Tes Alfa-Beta memasuki wilayah public pada akhir perang dunia ke-1, beberapa penyusun dan penerbit tes mengadaptasi tes tersebut untuk didistribusikan pada khalayak ramai. Diantara tes-tes yang diadaptasi dari tes Alfa-Beta, sub tes Beta terdiri atas lima sub skala performa (Cornell dan Coxe, 1934) yakni : (1) *Picture arrangement* atau merangkai gambar; (2) *Digit symbol*; (3) *Picture completion*

atau melengkapi gambar; (4) *Memory for designs* atau hafalan desain; dan (5) *Cube construction* atau konstruksi kubus.

Wechsler (Jakson 2003:42) mengadaptasi seperangkat sub tes performa dari Skala Poin Cornell-Coxe dan menggabungkannya dengan kelima sub-tes pada Skala Inteligensi Wechsler-Bellevue. Namun, Wechsler mengganti konstruksi kubus dengan balok susun dari Kohs (1923). Meskipun Skala Wechsler memiliki sejumlah fitur positif termasuk gagasan untuk mendasarkan perkiraan inteligensi atas sejumlah pengukuran verbal dan performa yang berbeda-beda serta pelaporan dengan bentuk profil, namun kontribusi Wechsler bukanlah sebagai yang pertama mengimplementasikan gagasan tersebut, bukan pula yang pertama mengembangkan materi tes orisinal, terlepas dari itu Wechsler lah yang mengombinasikan fitur-fitur tersebut dalam satu paket yang bermanfaat dan disesuaikan dengan penormaan berbasis usia. Sejak publikasi pertamanya, skala Wechsler telah menghasilkan beragam penelitian lanjutan yang penting berupa artikel dan data validitas bukan hanya mengenai property skala Wechsler tapi juga mengenai konsep inteligensi.

5. Ciri Khas Fitur Tes MAB II

Terlepas dari penerimaan dan penggunaan secara luas dari tes inteligensi individual, bentuk pengadministrasian individual memiliki kelemahan untuk digunakan di masa sekarang ini. Kelemahan pertama berkenaan dengan pengadministrasian secara individual, kelemahan kedua berhubungan dengan penyekoran judgmental, dan yang kelemahan ketiga berkaitan dengan penormaan. Dalam pelaksanaan tes inteligensi secara individual, pengadministrasian dan penyekoran harus dilakukan oleh tester profesional yang terlatih secara khusus. Hal tersebut tentunya memerlukan biaya yang lebih mahal dan cenderung mengalihkan profesional dari aktivitas lain yang lebih penting dan lebih menguntungkan secara keilmuan.

Pertimbangan kedua dalam pemilihan profil hasil pengukuran kemampuan intelektual umum terkait dengan pertanyaan mengenai penyekoran

terstruktur versus penyekoran judgmental. Penyekoran judgemental secara tradisional telah diterapkan dalam pengukuran berorientasi klinis terhadap inteligensi, namun penggunaannya memerlukan perhatian yang cermat seperti. Beberapa kekurangan penyekoran judgmental antara lain : (a) Memakan waktu lama dan dapat mengalihkan waktu untuk kegiatan profesional, (b) Memperkenalkan dimensi ketidakreliabelan pada penyekoran tes, dan (c) Memperkenalkan pada kesulitan dalam mengkonstruksi tes dan membakukan tes karena analisis item dan pemilihan item secara teliti memakan banyak biaya dan kurang reliabel dalam penyekoran tes secara judgemental. Penyekoran judgemental mengukur hasil tes memiliki masalah yang sangat serius yakni ketidakreliabelan penyekoran. Sejumlah studi (Ryan, Prifiteria dan Powers, 1983) mengimplikasikan bahwa para penyekor menunjukkan kesenjangan yang kentara dalam mengevaluasi skala Wechsler. Kesenjangan tersebut sebagian besar berupa variansi eror yang terdapat dalam penyekoran tidak menghilang sebagaimana fungsi dari tingkat keterdidikan (tingkat pendidikan) penyekor. Reliabilitas penyekoran tes merupakan hal yang perlu diperhatikan oleh professional yang peduli terhadap kualitas pelayanan mereka.

Kelemahan ketiga dari tes inteligensi yang diadministrasikan secara individual adalah masalah penormaan. Lebih sulit untuk mengambil sampel representatif dari responden tes individual dibandingkan dari responden tes kelompok.

Dalam rangka menindak lanjuti pertimbangan-pertimbangan tersebut serta mengintegrasikan keuntungan dari pendekatan pengukuran modern terhadap tes terstruktur yang terstandarisasi dengan format profil tes individu dari pengukuran kemampuan umum, maka disusunlah tes MAB pertama. Dalam rangka member informasi diagnostik, MAB-II telah dirancang menggunakan format yang memungkinkan administrasi tes secara kelompok dan terotomatisasi, penyekoran manual atau otomatis (menggunakan mesin atau komputer) serta acuan terhadap sistem penormaan tes terkini. Edisi pertama MAB menggunakan norma yang sama dengan WAIS-R, sementara edisi kedua MAB berdasarkan

pada penormaan Amerika Utara yang terbaru, penormaan tersebut berbasis data sensus Amerika Serikat dan sejumlah sampel responden dari Kanada.

6. Pengadministrasian Tes MAB II

Tes MAB-II merupakan tes yang terstandarisasi, memerlukan perhatian yang teliti terhadap instruksi dan alokasi waktu. Apabila pelaksanaan tes memenuhi prosedur administrasi standar maka hasil tes akan dapat diinterpretasi. Pengadministrasian dan penyekoran manual maupun komputerisasi tersedia untuk tes MAB-II.

Untuk pengadministrasian secara kelompok disarankan satu orang pengawas untuk 25 orang testi pertama dan satu orang proktor untuk 25 orang testi selanjutnya. Sementara untuk pengadministrasian tes secara individual tester atau proktor harus ada di ruangan tes selama tes berlangsung. Bab ini pada dasarnya membahas instruksi untuk pengadministrasian tes secara kelompok, namun dapat pula diaplikasikan pada pengadministrasian secara individual dengan sedikit modifikasi.

Untuk detail yang lebih spesifik mengenai pengadministrasian dan penyekoran secara komputerisasi, dapat digunakan program *SigmaSoft MAB-II for Windows Manual*. Program tersebut memfasilitasi pengadministrasian tes MAB-II secara komputerisasi pada semua bagian termasuk sub-tes performa, langsung dari komputer. *SigmaSoft MAB-II for Windows Manual* membuat pengadministrasian tes MAB-II menjadi lebih mudah dan menjamin prosedur standar administrasi tes terlaksana dengan baik. Setelah tester memilih sub-tes yang diinginkan, maka selanjutnya komputer mengambil alih pelaksanaan tes dari mulai menyajikan instruksi umum, latihan soal, contoh soal, sampai waktu tes. Piranti lunak *SigmaSoft MAB-II for Windows Manual* juga dirancang untuk memfasilitasi penyekoran sesingkat mungkin. Empat jenis laporan hasil tes dapat diproduksi yakni laporan data (*data report*), laporan dasar (*basic report*), laporan klinis (*clinical report*) dan laporan lanjutan (*extended report).* Silakan gunakan

SigmaSoft MAB-II for Windows Manual untuk mengetahui deskripsi lebih mendetail mengenai keempat jenis laporan tersebut.

a. Instrumen dan kelengkapan tes MAB II

MAB II terdiri dari dua bagian yaitu tes Verbal dan Tes performa yang masing masing dituangkan kedalam sebuah buku tes. Dalam rangka pengadministrasian tes kedua buku tersebut dapat diberikan secara bergantian baik verbal terlebih dahulu, ataupun tes performa, mengenai kelengkapan lain yang harus diperhatikan adalah kelengkapan lembar jawaban, kertas buram sebagai alat bantu untuk menghitung dan alat pengatur waktu sebagai patokan pelaksanaan tes.

b. Instruksi dalam Tes MAB II

Tester harus memiliki salinan dari seluruh materi tes dan menyiapkan diri untuk menjawab pertanyaan yang diajukan oleh testi. Beberapa pernyataan yang di cetak miring pada bagian ini harus dibaca dengan suara keras. Pernyataan yang tidak dicetak miring hanya anjuran bagi tester dan tidak usah dibaca dengan suara keras. Semua bagian sub-tes hanya diberi waktu pengerjaan selama 7 menit, tidak kurang, tidak lebih. Tester harus mengikuti batasan waktu tersebut dengan akurat, dan memastikan bahwa seluruh testi mengerjakan tes yang benar sesuai dengan waktunya.

7. Hal Lain yang Perlu Diperhatikan dalam Pelaksanaan Tes MAB II

a. Buku Tes

Untuk mengadministrasikan keseluruhan tes MAB-II, tersedia dua buku tes beserta lembar jawabannya, begitu pula format laporan MAB-II untuk penyekoran manual. Buku tes tes terdiri atas MAB-II Verbal dan MAB-II Performa, keduanya identik dengan buku tes pada MAB edisi pertama. Format laporan MAB-II mengalami revisi dari edisi sebelumnya, sehingga diperlukan bagan konversi untuk melakukan penyekoran manual (bagan konversi terlampir). Format laporan MAB edisi pertama tidak boleh digunakan untuk MAB-II. Dalam pengadministrasian tes, tester akan memerlukan *stop watch* dan testi/responden

memerlukan sebatang pensil 2B (tidak boleh HB) untuk menandai lembar jawaban. Selain itu, tester perlu menyediakan kertas stensil bagi responden untuk keperluan menghitung soal-soal pada bagian aritmatika sub-tes verbal.

b. Agenda Tes

Seluruh tes harus diberikan sesuai dengan urutan yang ada pada buku tes dan lembar jawaban. Pembacaan instruksi untuk tiap bagian tes memerlukan waktu sekitar dua- tiga menit. Seluruh bagian tes memiliki waktu pengerjaan selama tujuh menit. Penting untuk menjaga batas waktu ini dengan ketat dengan tujuan memperoleh penilaian yang akurat terhadap tingkat kemampuan testi.

Setengah dari waktu pelaksanaan tes MAB-II dialokasikan untuk sub-tes verbal sementara setengah lainnya untuk sub-tes performa. Waktu untuk melaksanakan kedua sub-tes tersebut sampai selesai adalah sekitar 50 menit. Tester dapat memilih untuk memberikan masing-masing sub-tes pada hari yang berbeda, atau dengan jeda yang singkat di antara kedua sub-tes.

c. Posisi tempat duduk testi

Tempat duduk testi dalam ruangan tes diatur dengan mengacu pada seting edukasional, yakni dimana testi duduk terpisah satu bangku dari testi lainnya. Pengaturan posisi tempat duduk bias saja berbeda pada pelaksanaan masing-masing sub-tes selama testi tetap terpisah sejauh satu bangku dari testi lainnya.

d. Pendistribusian materi tes

Apabila sejumlah testi di tes pada waktu yang sama, sebaiknya buku tes dan lembar jawaban di bagikan sebelum testi duduk atau memasuki ruangan tes. Apabila testi yang di tes hanya beberapa orang saja (misalnya kurang dari 5 orang) maka buku tes dan lembar jawaban dapat dibagikan setelah testi duduk. Apabila sub-tes verbal dan performa akan diadministrasikan dalam satu sesi, sub-tes verbal harus dilaksanakan lebih dulu, oleh karena itu buku tes verbal yang pertama kali dibagikan.

e. Keamanan Tes

Sangat penting untuk menjaga buku tes dan kunci jawaban MAB-II (baik asli maupun salinan) agar tidak jatuh ke tangan non-profesional, orang yang berpotensi menjadi testi maupun pihak lain yang tidak bertanggungjawab. Apabila pengadministrasian tes dilakukan secara kelompok, maka buku tes dan lembar jawaban harus di hitung sebelum dan sesudah pelaksanaan tes. Ketika tes selesai dilaksanakan, penting untuk mengumpulkan buku tes dan lembar jawaban dari seluruh testi sebelum mereka meninggalkan ruangan. Publikasi tes secara illegal maupun pemberian buku tes tanpa izin kepada calon testi tertentu, akan merugikan testi lain karena pelaksanaan tes menjadi tidak adil. Hal tersebut juga menyalahi kode etik praktik asesmen dan pengukuran.

8. Pengolahan Skor dan Interpretasi Tes MAB II

a. Penyekoran Manual

Penyekoran manual atau *hand scoring* memerlukan waktu kurang lebih 10 menit dan memerlukan seperangkat *template* penyekoran (kunci jawaban), lembar jawaban testi, format laporan hasil tes MAB-II dan tabel konversi skor. Format laporan hasil tes MAB-II menyajikan skala skor yang ekuivalen dengan skor mentah untuk digunakan dalam penghitungan skala skor verbal, performa maupun skor total. Tabel konversi mencakup skala verbal, skala performa, total skor dan skor standar yang ekuivalen dengan skor skala total untuk setiap kelompok umur yang berbeda. Berikut langkah-langkah yang harus ditempuh untuk mengkonversikan skor mentah menjadi skor skala juga untuk menggunakan tabel konversi dan format laporan hasil tes.

1. Periksa lembar jawaban tes verbal (warna hijau) dan lembar jawaban tes performa (warna ungu) dengan menggunakan template penyekoran atau kunci jawaban. Pastikan pemeriksa menuliskan total skor mentah pada masing-masing bagian sub-tes verbal maupun performa, kemudian menuliskan skor total di kolom bertuliskan skor mentah di format laporan hasil tes MAB-II.

2. Pengkonversian skor mentah menjadi skor T (dengan rata-rata 50 dan standar deviasi 10 dalam analisis yang terstandarisasi) dilakukan dengan cara melingkari skor mentah pada kolom yang terdapat pada bagian profil skala. Pengkonversian harus dilakukan pada kesepuluh bagian tes.
3. Skor skala pada masing-masing bagian tes kemudian dapat dibaca pada kolom paling kanan, kolom paling kiri, maupun kolom bagian tengah bagan profil skala pada baris horizontal yang sejajar dengan skor mentah yang dilingkari. Skor skala tersebut harus dicatat pada kolom sebelah kanan skor mentah di bawah tulisan "skor skala' pada baris yang diberi label sesuai bagian sub-tes verbal maupun performa.
4. Jumlahkan skor skala kesepuluh bagian sub-tes kemudian tuliskan jumlahnya pada ruang yang tersedia di format laporan hasil tes (apabila jumlah bagian tes yang diadministrasikan kurang dari lima maka penjumlahannya harus mengacu pada lampiran d 'memprorasikan skor sub-tes'). Jumlahkan skor sub-tes verbal dan performa untuk memperoleh skor total.
5. Dalam menghitung usia testi, tester harus menggunakan usia pada waktu tanggal kelahiran terakhir testi bukan membulatkannya. Perhatikan tabel konversi sesuai usia testi, setiap kategori usia memiliki konversi skor verbal, skor performa dan skor total yang berbeda-beda.
6. Mulailah pada kolom dengan nilai skor skala total paling rendah kemudian lihat terus ke bawah halaman, sampai menemukan nilai dari entri pada kolom berlabel Total skor skala yang sama dengan atau lebih besar dari total skor skala testi. Total skor skala untuk tes verbal, performa dan gabungan keduanya harus ditulis pada ruang yang tersedia di format laporan hasil tes. Apabila nilai skor skala testi sama dengan nilai pada tabel/kolom catat pula IQ dan persentil yang berada pada kolom sejajar skor skala testi. Apabila skor skala testi lebih kecil dari pada nilai pada tabel/kolom maka perhatikan nilai yang berada di bawahnya dan catat IQ dan persentil yang berada pada kolom tersebut. Perlu diingat bahwa dalam prosedur ini tidak dibolehkan adanya pembulatan atau prorasi. Baik IQ, skor standar, dan persentil harus setara

dengan nilai pada tabel/kolom yang sejajar atau berada dibawah nilai total skor skala testi.

7. Catatlah skor IQ dan persentil yang telah diperoleh pada format laporan hasil tes MAB-II pada masing-masing bagian verbal, performa dan skor total.

Apabila hanya dua, tiga, empat bagian sub-tes verbal atau hanya dua, tiga atau empat bagian sub-tes performa yang diadministrasikan, maka pemeriksa dapat memprorasikan skor bagian-bagian sub-tes tersebut dengan mengikuti petunjuk pada lampiran 'd'. Prorasi skor menyajikan dasar untuk mengidentifikasi nilai yang diharapkan dari total skor skala serta IQ verbal maupun performa yang tetap berdasarkan atas kelima bagian sub-tes meskipun bagian sub-tes yang diadministrasikan kurang dari lima.

Pihak pelaksana tes yang ingin menginterpretasikan pola skor sub-tes harus mengacu pada profil khusus dalam tabel skor standar yang ekivalen dengan total skor skala di lampiran 'b' disesuaikan dengan kategori usia testi. Profil tersebut akan memberikan dasar yang lebih akurat untuk interpretasi data, karena reliabilitas dari perbedaan antara poin-poin pada profil lebih rendah daripada skor skala individu, sehingga perbedaan substansial dapat ditemukan sebelum perbedaan yang sebenarnya diinterpretasi secara psikologis.

b. Penyekoran dengan Komputer

Penyekoran dengan computer dilakukan dengan menggunakan piranti lunak *SigmaSoft MAB-II for Windows*. Pemeriksa hanya perlu mengikuti petunjuk yang terdapat pada *SigmaSoft MAB-II for Windows Manual.*

c. Interpretasi Tes

Para pakar dan peneliti yang ingin mempelajari dan menggunakan data pada masing-masing sub-tes MAB-II harus mempertimbangkan poin mengenai perbedaan yang dapat diamati antara skor sub-tes testi pada tes MAB-II harus cukup besar agar reliabel secara statistik. Reliabilitas dari perbedaan sub-tes merupakan fungsi atas reliabilitas respektif dan kebebasan statistik dari skor-skor kesepuluh bagian tes. Sub-tes pada MAB-II secara substansial saling berkorelasi dikarenakan faktor general dari kemampuan intelektual umum. Meskipun sub-tes

pada MAB-II reliabel, namun keduanya memiliki variansi sama yang mempengaruhi interpretabilitas dari perbedaan antar sub-tes. Metode untuk mengevaluasi reliabilitas perbedaan tersebut disajikan oleh Crawford (1997).

9. Standarisasi dan Perangkat Psikometri dalam Tes MAB II

MAB-II melakukan penormaan kembali, oleh karena itu uji cobanya menggunakan sampel normatif. Sampel untuk penormaan terdiri atas sampel sistematis yang mencakup individu dari sembilan kelompok usia di Amerika Serikat dan Kanada. Hal tersebut dimaksudkan untuk menjamin partisipasi sampel dari beragam populasi umum dengan variasi ras dan etnisitas. Sampel mencakup individu berjenis kelamin pria dan wanita yang sama jumlahnya. Sembilan kelompok usia masing-masing terdiri atas 200 orang, dengan pengecualian terhadap dua kelompok usia paling tua yang masing-masing hanya terdiri atas 100 orang.

Tabel 2.3

Standarisasi tes MAB II di Amerika

SEX	N
Males	800
Females	800
RACE	**%**
White	87.4%
Non-white	12.6%
AGE GROUP	**N**
16-17	200
18-19	200
20-24	200
25-34	200
35-44	200
45-54	200
55-64	200
65-69	100
70-74	100
EDUCATION	**Years**
Mean Education Level	12.92
OCCUPATION	**%**
Managerial/Professional and specialty	50%
Technical, sales, service and labor	50%
REGION / COUNTRY	**%**
Northwest US and Western Canada	21.5
Southwest US	17.1
Midwest US	29.9
Southeast US	8.9
Northeast US and Eastern Canada	22.6
TOTAL	**N = 1600**

1. Norma Kelompok Usia

Rata-rata dan standar deviasi skor mentah masing-masing sub-tes dihitung untuk setiap kelompok usia. Karena rata-rata tingkat pendidikan tiga kelompok sampel (55-64, 65-69, dan 70-74) lebih tinggi dari rata-rata tingkat

pendidikan kelompok populasi, sehingga skor mentah dari ketiga kelompok ini perlu disesuaikan menggunakan modifikasi berdasarkan pertimbangan rata-rata edukasi populasi, rata-rata populasi sampel, serta korelasi antara skor sub-tes dan tingkat pendidikan. Rata-rata dan standar deviasi untuk setiap bagian tes kemudian diplot dan diperhalus menggunakan kriteria pangkat terkecil. Rata-rata dan standar deviasi yang dihasilkan dari proses tersebut kemudian digunakan untuk menghitung skor skala (skor T) dan persentil yang ekivalen dengan skor mentah. Skor skala ekivalen ini digunakan untuk membandingkan skor responden hanya dengan kelompok yang seusia. Skor ini tidak dapat digunakan untuk menkonversi skor mentah skor skala dalam penghitungan skor tes verbal, performa maupun skor total, karena penghitungan tersebut harus menggunakan bagan konversi skor skala.

2. Norma Konversi ke Skor Skala

Rata-rata dan standar deviasi skor mentah setiap bagian tes juga dihitung untuk seluruh kelompok sampel normatif. Sekali lagi, skor untuk tiga kelompok sampel usia yang paling tua disesuaikan dengan menggunakan modifikasi regresi. Rata-rata dan standar deviasi skor sub-tes seluruh kelompok sampel juga diplot dan diperhalus menggunakan criteria pangkat terkecil. Rata-rata dan standar deviasi yang dihasilkan dari penghitungan tersebut digunakan untuk memproduksi skor skala yang ekivalen dengan skor mentah yang terdapat dalam format laporan hasil MAB-II.

3. Norma Tes Skor Tes Verbal, Tes Performa dan Skor Total

Skor verbal, performa dan skor total dihitung untuk sampel pada masing-masing kelompok usia menggunakan rata-rata dan standar deviasi sebagaimana dijelaskan sebelumnya. Skor tersebut digunakan untuk menghitung rata-rata dan standar deviasi pada masing-masing kelompok usia, yang kemudian diperhalus dan dilaporkan seperti di atas dan digunakan untuk menghasilkan skor IQ, skor standar dan persentil.

10. Korelasi MAB II dengan Tes lain yang Sejenis

1. Korelasi MAB dengan WAIS-R

Korelasi antara MAB dan WAIS-R tersaji dalam table 4.8. Untuk tujuan perbandingan dalam table 2.4 juga disajikan ringkasan dari lima buah studi yang melaporkan korelasi antara WAIS dan WAIS-R. Perbandingan tersebut disajikan karena, tidak seperti perbandingan WAIS dan MAB, WAIS dan WAIS-R memiliki kesamaan metode pengadministrasian, format item, dan mayoritas item soal yang identik. Korelasi antara sub-tes pada MAB dan WAIS-R berada pada rentang korelasi antara sub-tes pada WAIS dan WAIS-R, dengan pengecualian tunggal pada sub-tes Spatial dan Block Design. Korelasi skor total pada MAB dan WAIS-R memiliki nilai sebesar 0,91. Nilai tersebut hampir sama dengan nilai rata-rata korelasi skor total WAIS dan WAIS-R yakni 0,93.

Tabel 2.4

Korelasi dengan WAIS-R

	MAB/WAIS-R[a]	WAIS/WAIS-R[b]	
Subtest	r	r avg.	range
Information	.82	.87	.81 - .91
Comprehension	.73	.80	.52 - .93
Arithmetic	.89	.82	.27 - .93
Similarities	.66	.78	.33 - .90
Vocabulary	.89	.86	.26 - .96
Digit Symbol	.45	.84	.29 - .98
Picture Completion	.87	.74	.30 - .90
Spatial/Block Design	.44	.88	.74 - .97
Picture Arrangement	.86	.62	.15 - .84
Object Assembly	.65	.80	.14 - .98
Verbal Scale	.94	.91	.73 - .96
Performance Scale	.79	.89	.76 - .97
Full Scale	.91	.93	.85 - .98

Note. [a] From Jackson (1984), N = 145, mixed sample; correlations computed using scaled scores (for subtests) and sums of scaled scores (for scale totals).

[b] Correlations reported by Wechsler (1981), Lippold & Claiborn (1983), Mishra & Brown (1983), Rabourn (1983), and Smith (1983) were averaged using Fisher's *r* to *z* transformation.

2. Perbandingan MAB dengan *Advanced Progressive Matrices* (APM) dari Raven

Advanced Progressive Matrices (APM) yang dikembangkan oleh Raven (1960) merupakan pengukuran non-verbal yang banyak dipergunakan untuk mengukur kemampuan intelektual umum. Tabel 2.5 menyajikan korelasi MAB dengan APM, berdasarkan sampel yang diambil dari mahasiswa University of Calofornia di Berkeley, Amerika Serikat (Kranzler dan Jensen, 1991). Meskipun magnitudo korelasi tersebut meningkat secara substansial dikarenakan pembatasan rentang pada populasi ini, pola korelasi tidak sama dari sub-tes satu pada sub-tes lainnya. Korelasi tertinggi dihasilkan oleh sub-tes performa pada MAB yakni *Object Assembly dan Spatial*, sementara pada sub-tes verbal korelasi tertinggi dihasilkan oleh *Arithmetic* yang merupakan suatu pengukuran kemampuan nalar. Selain *Arithmetic* semua bagian tes verbal menunjukkan korelasi yang lebih rendah. Dengan demikian, jelas bahwa tes APM dari Raven lebih baik dalam mengukur komponen cair (*fluid components*) dari kemampuan intelektual umum yang terdapat pada MAB sub-tes performa daripada mengukur komponen terkristalisasi (*crystallized components*) yang terdapat pada MAB sub-tes verbal.

Tabel 2.5

Korelasi antara tes MAB dengan tes APM dari Raven (N = 103)

Subtest	Correlation
Verbal	
Information	.19
Comprehension	.23
Arithmetic	.46
Similarities	.13
Vocabulary	.14
Performance	
Digit Symbol	.25
Picture Completion	.30
Spatial	.44
Picture Arrangement	.31
Object Assembly	.56

Note. Adapted from Kranzler and Jensen (1991) with corrections from J. H. Kranzler, personal communication to Michael C. Ashton, April 15, 1998