

Perbedaan Waktu Pemberian Jawaban dan Hasil Tes Inteligensi Ditinjau dari Perbedaan Lembar Jawaban

*Ridwan Saptoto*¹

Fakultas Psikologi
Universitas Gadjah Mada

Abstract

Many intelligence test administered designed as paper and pencil test. But efficiency demand in this time make it administered using computer answer sheet, and testie forced to give answer by blackening answer dots using 2B computer pencil in answer sheet. This method reduce testie time to think the right answer, and may influence intelligence test result. This study aimed to compare time to fill answer sheet and intelligence test results, between testie who did test with paper and pencil answer sheet and testie who did test with computer answer sheet. Experiment method choseed in this study. Simple design with repeated sample used in first experiment step. Prior, subjects were assigned to fill answer sheet by streak. And then they assigned to give answer by blackening answer dots. Data analysis show that there were significant difference between first and second assignment ($F = 192.192, p < 0.05$). Two independent group design used in the second experiment step. Subjects were randomly assigned to join control group or experiment group. Subjects in control group did test using paper and pencil answer sheet, while subjects in experiment group did test using computer answer sheet. Data analysis show that there were significant difference in intelligence level between control group and experiment group in CFIT A ($t = 1.681, p < 0.05$; one tailed), CFIT B ($t = 2.289, p < 0.05$; one tailed), and CFIT A and B ($t = 1.819, p < 0.05$; one tailed). Control group consistently higher in intelligence test results than experiment group.

Keywords: CFIT, computer answer sheet, intelligence test

Pengukuran atribut psikologis menjadi salah satu topik yang menarik untuk diteliti. Para ahli psikologi semenjak dahulu tertantang untuk mengembangkan berbagai alat ukur atribut-atribut psikologis. Hal inilah yang kemudian melahirkan berbagai model alat ukur dalam ilmu psikologi.

Perkembangan zaman dan teknologi komputer ternyata juga mempengaruhi metode dan prosedur pengukuran. Saat ini

komputer telah digunakan dalam proses pemeriksaan dan pengetesan. Subjek tidak lagi mengerjakan tes pada kertas lembar jawaban, namun langsung mengerjakannya pada komputer. Hasil penelitian yang dilakukan Kumar dan Helgeson (1995) menunjukkan bahwa penggunaan komputer pada proses pemeriksaan atau pengetesan mampu meningkatkan kualitas dan efisiensi perekaman, penyimpanan, dan analisis data. Fisteus, Pardo, dan Garcia (2012) juga mengemukakan bahwa penggunaan komputer akan mengurangi

¹ Korespondensi mengenai artikel ini dapat dilakukan melalui: rsaptoto@gmail.com

waktu dalam koreksi dan penilaian, serta meningkatkan reliabilitas hasil penelitian dibandingkan metode konvensional.

Perkembangan lain dari teknologi komputer yaitu penggunaan lembar jawaban komputer dalam pengetesan. Metode ini untuk mempercepat proses koreksi hasil tes dengan menggunakan program komputer, terutama pada tes yang dilakukan secara klasikal dan melibatkan ratusan atau bahkan ribuan *testi*.

Metode tersebut mengharuskan *testi* untuk menghitamkan bulatan alternatif jawaban yang dipilihnya. Lembar jawaban tersebut kemudian akan di-*scan* oleh sebuah alat, sehingga hasil tes dapat diketahui dengan singkat. Metode ini sangat menghemat waktu dan biaya dalam sebuah pemeriksaan yang menggunakan tes psikologis. Tenaga *testir* untuk memeriksa dan menghitung hasil pengetesan juga dapat dihemat.

Sebuah perusahaan bernama SMR Indonesia (2012) di dalam *web site*-nya secara jelas menyebutkan bahwa mereka menyediakan program komputer untuk membaca hasil psikotes yang dikerjakan pada lembar jawaban komputer. Jenis-jenis psikotes yang hasilnya dapat dibaca oleh program tersebut adalah CFIT, abstrak, relasi ruang, mekanik, skolastik, dan klerikal.

Peneliti menemukan bahwa selama ini belum ada peneliti lain yang meneliti pengaruh penggunaan lembar jawaban komputer terhadap hasil tes, terutama tes yang pengerjaannya dibatasi waktu secara ketat. Peneliti juga menemukan belum ada jurnal yang membahas pengaruh penggunaan lembar jawaban komputer terhadap hasil tes.

Peneliti hanya menemukan hasil penelitian mengenai penggunaan aplikasi komputer untuk melakukan pengetesan

secara virtual yang dilakukan oleh Kumar dan Helgeson (1995). Namun pengetesan tersebut dilakukan dengan membuat aplikasi di komputer, dan subjek penelitian mengerjakan tes secara langsung di komputer. Para subjek tidak mengerjakan tes dengan cara menghitamkan bulatan alternatif jawaban yang dipilihnya. Oleh karena itu, peneliti ingin mengkaji hal ini secara lebih lanjut di dalam penelitian ini. Peneliti ingin mengetahui bagaimana dampak penggunaan lembar jawaban komputer pada tes psikologis.

Anastasi dan Urbina (2007) mendefinisikan tes psikologis sebagai alat ukur yang objektif dan distandardisasikan atau dibakukan pengukurannya berdasarkan contoh perilaku tertentu. Hasil dari tes psikologis akan memberikan informasi mengenai aspek-aspek psikologis dari individu yang menjadi peserta tes. Beberapa jenis tes psikologis yaitu tes inteligensi, tes bakat, dan tes kepribadian.

Berbagai ahli mengemukakan pendapatnya mengenai inteligensi. Terman mendefinisikan inteligensi sebagai kemampuan untuk berpikir abstrak (Azwar, 2010). Flynn (dalam Baron, 1996) memberikan definisi yang serupa. Menurutnya, inteligensi merupakan kemampuan seseorang untuk berfikir secara abstrak dan kesiapannya untuk belajar dari pengalaman.

Binet dan Simon (dalam Azwar, 2010) selanjutnya mengatakan bahwa inteligensi terdiri dari tiga komponen. Ketiga komponen tersebut adalah kemampuan untuk mengarahkan pikiran atau tindakan, kemampuan untuk mengubah arah tindakan bila tindakan tersebut telah dilakukan, dan kemampuan untuk mengkritik atau mengoreksi diri sendiri.

Wechsler (dalam Azwar, 2010) mengemukakan bahwa inteligensi merupakan kumpulan atau totalitas kemampuan seseorang. Totalitas tersebut mencakup

kemampuan untuk bertindak dengan tujuan tertentu, kemampuan berpikir secara rasional, dan kemampuan untuk menghadapi lingkungannya secara efektif.

Cattell (1971) berpendapat bahwa inteligensi atau *general (g) intelligence* terdiri dari dua faktor, yaitu *fluid intelligence* dan *crystalized intelligence*. *Fluid intelligence* merupakan kapasitas seseorang untuk berpikir secara logis dan menyelesaikan masalah dalam situasi yang baru, dan mandiri dalam mengumpulkan pengetahuan yang dibutuhkan. Termasuk di dalamnya yaitu kemampuan untuk menganalisis permasalahan-permasalahan baru, mengidentifikasi pola-pola dan hubungan-hubungan antar masalah, serta membuat perhitungan secara logis. Termasuk di dalam *fluid intelligence* yaitu kemampuan untuk menggunakan gaya berpikir deduktif maupun induktif.

Cattell (1971) selanjutnya menyatakan bahwa *crystalized intelligence* merupakan kemampuan seseorang untuk menggunakan keterampilan, pengetahuan, dan pengalamannya. Hal ini membuat *crystalized intelligence* pada umumnya akan berkembang sesuai dengan penambahan usia seseorang. *Crystalized intelligence* diidentifikasi dari keluasan dan kedalaman pengetahuan umum yang dimiliki seseorang, kosa kata, dan kemampuan untuk berpikir menggunakan kata-kata maupun angka. *Crystalized intelligence* merupakan produk dari pendidikan dan pengalaman seseorang dalam interaksinya dengan *fluid intelligence*.

Berdasarkan pandangan dari berbagai ahli di atas, maka peneliti menggunakan definisi inteligensi yang dikemukakan oleh Cattell (1971). Hal ini karena definisi tersebut sudah sesuai dengan penelitian yang akan dilaksanakan.

Para ahli mengembangkan berbagai pengukuran inteligensi atau tes inteligensi.

Tes inteligensi secara umum dapat didefinisikan sebagai tes yang mengungkap kemampuan mental seseorang (Azwar, 2010). Berbagai jenis tes inteligensi dikembangkan sesuai dengan definisi mengenai inteligensi yang diacu oleh para pembuat tes tersebut.

Binet dan Simon (dalam Azwar, 2010) pada tahun 1904 mulai mengembangkan skala inteligensi pertama yang ditujukan untuk membedakan anak-anak lemah mental dengan anak normal. Wechsler (dalam Azwar, 2010) juga mengembangkan tes inteligensi. Versi pertama tes inteligensi dari Wechsler diterbitkan pada tahun 1939. Tes inteligensi tersebut dirancang secara khusus untuk mengukur inteligensi orang dewasa.

Cattell (1949) juga mengembangkan tes inteligensi berdasarkan pandangannya sendiri mengenai inteligensi. Tes yang dikembangkan oleh Cattell disebut sebagai *Culture Fair Intelligence Test*, atau disingkat CFIT.

Cattell dan Cattell (1973) menyebutkan bahwa CFIT terdiri dari tiga jenis tes atau skala, yaitu skala 1, skala 2, dan skala 3. Skala 1 dipergunakan untuk mengukur inteligensi kecerdasan anak-anak berumur empat sampai dengan delapan tahun dan orang-orang yang lebih tua namun memiliki kesulitan belajar. Skala 2 dipergunakan untuk mengukur inteligensi anak-anak yang berusia delapan sampai dengan empatbelas tahun dan orang dewasa yang memiliki kecerdasan normal. Skala 3 dipergunakan untuk mengukur inteligensi orang berusia empatbelas tahun ke atas dan orang dewasa yang memiliki taraf kecerdasan superior. Skala CFIT 2 dan skala CFIT 3 memiliki bentuk paralelnya, yaitu *form A* dan *form B*. Hal ini membuat skala CFIT 2 terdiri dari CFIT 2A dan 2B, sedangkan skala CFIT 3 terdiri dari CFIT 3A dan 3B.

Cattell dan Cattell (1973) selanjutnya menyebutkan bahwa skala CFIT 3A dan 3B terdiri dari empat subtes. Subtes-subtes tersebut yaitu *series*, *classification*, *matrices*, dan *conditions* atau *topology*. Masing-masing tes harus dikerjakan dalam waktu yang telah ditetapkan. Peserta tes atau *testi* dipebolehkan mengerjakan tes setelah mendapatkan aba-aba dari *testir*. *Testi* selanjutnya harus berhenti mengerjakan tes pada saat *testir* memerintahkan untuk berhenti. Cara menjawab tes ini yaitu dengan cara memberikan tanda centang pada lembar jawaban CFIT.

Cattell (1961) menyebutkan bahwa pengerjaan skala 1 sebenarnya tidak dibatasi oleh waktu. *Testi* umumnya dapat menyelesaikan skala 1 dalam waktu 22 menit. Skala 2 dan skala 3 pada sisi lain pengerjaannya dibatasi. Kedua skala tersebut hanya diberi waktu pengerjaan selama 12,5 menit. Rinciannya yaitu *series* dikerjakan selama tiga menit, *classification* dikerjakan selama empat menit, *matrices* dikerjakan selama tiga menit, dan *conditions* atau *topology* dikerjakan selama 2,5 menit. Batasan waktu di atas digunakan baik untuk *form A* maupun *form B*.

Cattell (1961) juga mengatakan bahwa CFIT 3A hendaknya dikerjakan lebih dahulu daripada CFIT 3B. Masing-masing tes akan menghasilkan *raw score* yang kemudian diubah menjadi angka inteligensi dengan menggunakan norma yang sudah ada. *Raw score* dari kedua skala juga dapat digabungkan, sehingga menghasilkan *raw score* gabungan. *Raw score* gabungan tersebut selanjutnya diubah menjadi angka inteligensi total dengan menggunakan norma yang sudah tersedia.

Berdasarkan uraian di atas maka dapat disimpulkan bahwa CFIT terdiri dari tiga jenis tes atau skala, yaitu skala 1, skala 2, dan skala 3. Skala CFIT 2 dan skala CFIT 3 memiliki bentuk pararelnya, yaitu *form A*

dan *form B*. skala CFIT 3A dan 3B terdiri dari empat subtes. Subtes-subtes tersebut yaitu *series*, *classification*, *matrices*, dan *conditions* atau *topology*. Masing-masing tes harus dikerjakan dalam waktu yang telah ditetapkan.

CFIT selanjutnya disebut sebagai tes yang *culture fair* atau adil budaya, karena CFIT relatif bebas dari pengaruh tingkat pendidikan dan budaya dari *testi* (Cattell dan Cattell, 1973). Hal ini membuat CFIT relatif lebih mudah digunakan di berbagai negara yang memiliki budaya berbeda-beda tanpa harus melakukan adaptasi budaya terlebih dahulu. Kondisi ini menjadi kelebihan yang dimiliki oleh CFIT. Kelebihan tersebut tidak dimiliki oleh tes inteligensi yang dikemukakan oleh Wechsler.

Kelebihan yang dimiliki CFIT membuat para ahli masih terus menggunakan tes tersebut di dalam penelitiannya. Austin, Hofer, Deary dan Eber (2000) menggunakan CFIT untuk meneliti interaksi antara kepribadian dengan inteligensi. Hasil penelitiannya antara lain menemukan bahwa peningkatan inteligensi berbanding positif dengan peningkatan kecenderungan kecemasan (*neuroticism*) seseorang.

Gupta (1991) juga meneliti pengaruh kepribadian terhadap inteligensi. Variabel independen lain yang digunakannya adalah waktu pengetesan. Hasil penelitiannya menunjukkan bahwa kepribadian ternyata tidak mempengaruhi hasil tes CFIT. Hasil tes CFIT justru dipengaruhi oleh waktu pengerjaannya. Subjek yang mengerjakan CFIT di pagi hari ternyata mendapatkan skor rata-rata yang lebih tinggi dari mereka yang mengerjakan CFIT di siang atau sore hari.

Furlow, Armijo-Prewitt, Gangestad, dan Thornhill (1997) kemudian meneliti hubungan ketidaksimetrisan tubuh de-

ngan inteligensi. Mereka menggunakan CFIT untuk mengukur inteligensi. Hasil penelitian menunjukkan bahwa ketidak simetrisan tubuh ternyata memiliki hubungan negatif dengan inteligensi. Hal ini membuat mereka kemudian menyimpulkan bahwa meningkatkan kualitas lingkungan perkembangan seseorang akan meningkatkan inteligensinya.

Colom dan Garcia-Lopez (2002) selanjutnya meneliti pengaruh perbedaan jenis kelamin terhadap hasil tes inteligensi dari tiga jenis tes inteligensi. Hasil penelitiannya menunjukkan bahwa hanya CFIT yang menunjukkan tidak ada perbedaan tingkat inteligensi antara laki-laki dan wanita.

Kelebihan tersebut membuat Colom dan Garcia-Lopez (2003) kembali menggunakan CFIT untuk meneliti kecenderungan kenaikan inteligensi dari siswa sekolah menengah atas dan lulusan sekolah menengah atas di Spanyol. Hasil penelitiannya menunjukkan adanya kecenderungan kenaikan inteligensi secara umum dari kedua jenis kelamin subjek penelitian tersebut.

Rummsayer dan Brandler (2002) juga menggunakan CFIT untuk meneliti otak. Mereka meneliti hubungan antara inteligensi dengan indikator psikofisik dari resolusi sementara di dalam otak. Hasil penelitian mereka menunjukkan bahwa mekanisme otak secara spesifik terlibat di dalam diskriminasi interval pemrosesan informasi yang cepat pada mereka yang memiliki inteligensi lebih tinggi. Hal ini menunjukkan bahwa mekanisme tersebut dapat dipergunakan sebagai indikator inteligensi atau kecerdasan yang cukup sensitif.

Berdasarkan pendapat para ahli di atas, maka peneliti menggunakan CFIT untuk mengukur inteligensi di dalam penelitian ini. Hal ini karena CFIT relatif

adil budaya dan dapat dipergunakan tanpa harus diadaptasi terlebih dahulu. CFIT juga masih dipercaya oleh para ahli, sehingga mereka masih menggunakan CFIT untuk mengukur inteligensi di dalam penelitian-penelitiannya.

Tes CFIT pada saat ini sudah dapat dikerjakan dengan menggunakan lembar jawaban komputer (SMR Indonesia, 2012). Penggunaan jenis lembar jawaban tersebut membuat *testi* harus menghitamkan bulatan pada alternatif jawaban yang dipilihnya. Hal ini tentunya mengurangi waktu *testi* untuk mengerjakan soal, sehingga berpotensi mempengaruhi hasil tes inteligensi.

Tes CFIT sendiri menurut Cattell dan Cattell (1973) sebenarnya cukup dikerjakan dengan cara memberikan tanda centang pada lembar jawaban CFIT. Penggunaan lembar jawaban komputer membuat waktu yang dapat dipergunakan *testi* untuk berpikir dan menyelesaikan masalah menjadi berkurang, karena *testi* juga harus meluangkan waktu untuk menghitamkan bulatan alternatif jawaban yang dipilihnya. Kondisi ini berpotensi mempengaruhi hasil tes inteligensi.

Proses menghitamkan bulatan pada alternatif jawaban membutuhkan tambahan waktu. *Testir* sendiri pada umumnya ketika pelaksanaan tes psikologis tidak memberikan waktu tambahan, atau memberikan waktu khusus bagi *testi* untuk menghitamkan alternatif jawaban yang dipilihnya. *Testir* tetap berpedoman dan menggunakan standar waktu yang sama dengan cara pemberian jawaban yang dicoret.

Peneliti ketika melakukan asesmen di lapangan juga melihat bahwa subjek asesmen membutuhkan waktu yang lebih lama untuk memberikan jawaban pada lembar jawaban komputer. Waktu yang dibutuhkan untuk mengerjakan sebuah

soal sampai dengan selesai menghitamkan bulatan pada lembar jawaban komputer lebih lama, dibandingkan jika subjek asesmen mengerjakan soal pada lembar jawaban tradisional. Kondisi tersebut juga terjadi pada saat pemberian berbagai tes psikologis. Hal ini berpotensi mempengaruhi hasil tes psikologis.

Berdasarkan uraian-uraian di atas, maka peneliti mengajukan empat hipotesis penelitian. Hipotesis-hipotesis tersebut yaitu: (1) Terdapat perbedaan penggunaan waktu pengisian lembar jawaban ditinjau dari perbedaan penggunaan jenis lembar jawaban. Waktu yang dibutuhkan subjek untuk mengisi lembar jawaban manual lebih sedikit daripada waktu yang dibutuhkan subjek untuk mengisi lembar jawaban komputer. (2) Terdapat perbedaan hasil tes CFIT A ditinjau dari perbedaan penggunaan jenis lembar jawaban. Subjek yang menggunakan lembar jawaban manual mendapatkan hasil tes CFIT A yang lebih tinggi daripada subjek yang menggunakan lembar jawaban komputer. (3) Terdapat perbedaan hasil tes CFIT B ditinjau dari perbedaan penggunaan jenis lembar jawaban. Subjek yang menggunakan lembar jawaban manual mendapatkan hasil tes CFIT B yang lebih tinggi daripada subjek yang menggunakan lembar jawaban komputer. (4) Terdapat perbedaan hasil tes CFIT A dan B ditinjau dari perbedaan penggunaan jenis lembar jawaban. Subjek yang menggunakan lembar jawaban manual mendapatkan hasil tes CFIT A dan B yang lebih tinggi daripada subjek yang menggunakan lembar jawaban komputer.

Metode

Subjek penelitian

Subjek penelitian adalah para mahasiswa Fakultas "X" Universitas Gadjah Mada yang mendaftarkan diri dan berse-

dia mengikuti prosedur eksperimen yang telah disusun.

Instrumen penelitian

Data pertama, yaitu waktu pengisian lembar jawaban, akan diukur dengan menggunakan pengukur waktu atau *stopwatch*. *Stopwatch* yang digunakan memiliki ketelitian pengukuran sampai dengan seperseratus detik. *Stopwatch* tersebut sebelum digunakan dikalibrasikan dulu, sehingga setiap *stopwatch* dipastikan memiliki keakuratan yang sama.

Data penelitian kedua berupa inteligensi. Data inteligensi diambil dengan menggunakan CFIT 3A dan B yang disusun oleh Cattell dan Cattell (1973). Tes ini mengukur *general (g) intelligence* terdiri yang terdiri dari dua faktor, yaitu *fluid intelligence* dan *crystalized intelligence*. Goldstein dan Hersen (2000) mengungkapkan bahwa korelasi CFIT dengan tes-tes inteligensi yang lain menghasilkan koefisien korelasi sebesar 0,5 sampai dengan 0,7. Hal ini menunjukkan bahwa CFIT terbukti valid untuk mengukur inteligensi. Azwar (2009) mengemukakan bahwa sebuah tes yang valid berarti bahwa tes tersebut mampu menjalankan fungsi ukurnya sesuai maksud dikenakannya tes tersebut.

Goldstein dan Hersen (2000) selanjutnya mengungkapkan bahwa CFIT skala 3 memiliki koefisien konsistensi internal sebesar 0,85. Reliabilitas *test-retest* CFIT skala 3 adalah 0,82. Hal ini menunjukkan bahwa CFIT skala 3 terbukti reliabel dalam mengukur inteligensi. Azwar (2009) mengemukakan bahwa reliabilitas adalah keterandalan atau keajegan sebuah tes. Tes yang reliabel berarti bahwa tes tersebut memberikan hasil yang relatif sama dalam beberapa kali pengukuran yang melibatkan kelompok subjek yang sama pula, jika

aspek yang diukur dalam diri subjek memang tidak berubah.

Prosedur

Rekrutmen subjek penelitian dilakukan dengan cara memasang poster rekrutmen di lingkungan Fakultas "X" Universitas Gadjah Mada. Mahasiswa yang tertarik kemudian mendaftar menjadi subjek penelitian kepada asisten peneliti.

Mereka kemudian dikumpulkan untuk mengikuti eksperimen tahap pertama. Mereka pada awalnya diminta untuk mencoret alternatif jawaban pada lembar jawaban manual. Satu nomor hanya berisi satu alternatif jawaban yang sudah ditentukan. Penentuan alternatif jawaban tersebut dilakukan dengan menggunakan metode random. Asisten peneliti menghitung waktu yang diperlukan setiap subjek untuk mencoret semua alternatif jawaban yang sudah ditentukan.

Asisten peneliti kemudian memberikan lembar jawaban komputer kepada subjek. Mereka kini diminta untuk menghitamkan bulatan pada alternatif jawaban yang sudah ditentukan. Alternatif jawaban pada setiap nomor sama dengan alternatif jawaban pada langkah sebelumnya. Asisten peneliti kembali menghitung waktu yang diperlukan setiap subjek untuk menghitamkan bulatan semua alternatif jawaban yang sudah ditentukan.

Semua subjek selanjutnya dikelompokkan secara random ke dalam kelompok kontrol atau kelompok eksperimen untuk mengikuti eksperimen tahap kedua. Kelompok kontrol pada tahap pertama mengerjakan CFIT 3A dengan cara mencoret alternatif jawaban yang dipilihnya pada lembar jawaban. Mereka selanjutnya pada tahap kedua mengerjakan CFIT 3B, juga dengan cara mencoret alternatif jawaban yang dipilihnya pada lembar jawaban.

Kelompok eksperimen pada tahap pertama mengerjakan CFIT 3A dengan cara membulatkan hitam alternatif jawaban yang dipilihnya pada lembar jawaban. Dan pada tahap kedua mengerjakan CFIT 3B, juga dengan cara membulatkan hitam alternatif jawaban yang dipilihnya pada lembar jawaban.

Asisten peneliti menghitung jumlah jawaban benar dari setiap subjek kelompok kontrol maupun kelompok eksperimen yang telah mengerjakan skala CFIT 3A dan CFIT 3B. Jumlah jawaban benar tersebut disebut *raw score*, yang kemudian diubah menjadi angka inteligensi dengan menggunakan norma CFIT yang berlaku.

Pengambilan data dilakukan pada hari Jumat, 1 Juni 2012. Subjek yang mendaftar untuk mengikuti eksperimen berjumlah 46 (empatpuluh enam) mahasiswa. Asisten peneliti kemudian menempatkannya secara random ke dalam kelompok kontrol atau kelompok eksperimen. Pembagian dilakukan secara imbang jumlahnya, sehingga masing-masing kelompok terdiri dari 23 (duapuluh tiga) mahasiswa. Namun dalam kenyataannya, tidak semua subjek menghadiri pelaksanaan eksperimen. Hal ini membuat jumlah subjek menjadi berkurang. Peneliti mencatat kelompok kontrol maupun kelompok eksperimen pada akhirnya masing-masing hanya terdiri dari 13 (tigabelas) mahasiswa. Oleh karena itu, peneliti melakukan pengambilan data tahap kedua. Yang dilakukan pada Hari Kamis, 6 September 2012.

Subjek yang mendaftar untuk mengikuti eksperimen berjumlah 46 (empatpuluh enam) mahasiswa. Asisten peneliti juga menempatkan subjek secara random ke dalam kelompok kontrol atau kelompok eksperimen. Pembagian dilakukan secara imbang jumlahnya, sehingga masing-masing kelompok terdiri dari 23

(duapuluh tiga) mahasiswa. Namun dalam kenyataannya, kembali tidak semua subjek menghadiri pelaksanaan eksperimen. Hal ini membuat jumlah subjek menjadi berkurang. Peneliti mencatat bahwa subjek kelompok kontrol yang menghadiri pelaksanaan eksperimen berjumlah 22 (duapuluh dua) mahasiswa, sedangkan subjek kelompok eksperimen yang menghadiri pelaksanaan eksperimen berjumlah 12 (duabelas) mahasiswa. Hal ini menyebabkan secara keseluruhan data kelompok eksperimen yang dapat dianalisis menjadi 25 (duapuluh lima) set, sedangkan data kelompok kontrol yang dapat dianalisis menjadi 35 (tigapuluh lima) set.

Desain penelitian yang dipergunakan dalam penelitian ini adalah desain eksperimen. Penelitian ini sendiri terdiri dari dua tahap.

Penelitian tahap pertama menggunakan desain eksperimen yang disebut sebagai *simple design* (Myers, 1987). Desain ini memiliki beberapa bentuk. Bentuk yang digunakan peneliti dalam penelitian ini yaitu *simple design with repeated measure*. Masing-masing subjek penelitian akan dikenai dua kali pengukuran. Peneliti melakukan pengamatan terhadap semua pengukuran yang dilakukan. Gambar 1 menunjukkan desain penelitian tahap pertama yang digunakan.

KE	X_1O_1	X_2O_2
----	----------	----------

Gambar 1. Desain Penelitian Tahap Pertama

Keterangan:

KE : kelompok eksperimen

X_1O_1 : subjek mencoret alternatif jawaban, dan peneliti menghitung waktu yang diperlukan untuk memilih dan mencoret alternatif jawaban sesuai nomor soal pada lembar jawaban

X_2O_2 : subjek menghitamkan bulatan alternatif jawaban dan peneliti menghitung waktu

yang diperlukan untuk memilih dan menghitamkan bulatan alternatif jawaban sesuai nomor soal pada lembar jawaban

Peneliti selanjutnya pada tahap eksperimen yang kedua menggunakan *two independent group design* (Myers, 1987). *Two independent group design* merupakan desain eksperimen yang menggunakan dua kelompok. Masing-masing kelompok terdiri dari subjek penelitian yang berbeda. Kelompok pertama disebut sebagai kelompok kontrol (selanjutnya disingkat KK), dan kelompok kedua disebut sebagai kelompok eksperimen (selanjutnya disingkat KE). Masing-masing kelompok mendapatkan perlakuan yang berbeda.

Subjek dikelompokkan ke dalam kelompok kontrol atau kelompok eksperimen secara acak, atau disebut *random assignment*. Pembagian secara acak memungkinkan setiap subjek memiliki kesempatan yang sama untuk ditempatkan dalam salah satu kondisi perlakuan (Myers, 1987). Hal ini membuat kelompok berada dalam kondisi yang setara sebelum eksperimen dilakukan. Gambar 2 menunjukkan desain penelitian tahap kedua.

	KK	X_1O_1	X_2O_2
R			
	KE	X_3O_3	X_4O_4

Gambar 2. Desain Penelitian Tahap Kedua

Keterangan:

R : random

KK : kelompok kontrol

KE : kelompok eksperimen

X_1O_1 : subjek mengerjakan CFIT 3A dengan mencoret alternatif jawaban dan pengambilan data inteligensi oleh peneliti

X_2O_2 : subjek mengerjakan CFIT 3B dengan menghitamkan bulatan alternatif jawaban dan pengambilan data oleh peneliti

- X₃O₃ : subjek mengerjakan CFIT 3A dengan menghitung bulatan alternatif jawaban dan pengambilan data oleh peneliti
- X₄O₄ : subjek mengerjakan CFIT 3A dengan mencoret alternatif jawaban dan pengambilan data inteligensi oleh peneliti

Hipotesis-hipotesis penelitian yang diajukan dalam penelitian ini diuji dengan menggunakan metode statistik. Metode analisis data yang digunakan untuk menguji hipotesis penelitian pertama adalah *paired-sample t-test*. Selanjutnya metode analisis data yang digunakan untuk menguji hipotesis penelitian kedua, ketiga, dan keempat adalah *independent sample t-test*.

Pengujian statistik di atas akan dilakukan dengan menggunakan bantuan komputer. Program atau *software* yang digunakan adalah *Statistical Product and Service Solution (SPSS) 17.0 for Windows*. Taraf signifikansi yang digunakan sebesar 5%.

Hasil dan Diskusi

Paired-sample t-test antara waktu pengisian lembar jawaban manual dan lembar jawaban komputer menghasilkan ($p < 0.05$). Hipotesis penelitian pertama diterima. Terdapat perbedaan penggunaan waktu pengisian lembar jawaban ditinjau dari perbedaan penggunaan jenis lembar jawaban. Waktu yang dibutuhkan subjek untuk mengisi lembar jawaban manual lebih sedikit daripada waktu yang dibutuhkan subjek untuk mengisi lembar jawaban komputer.

Independent sample t-test terhadap angka inteligensi hasil skala CFIT 3A pada kelompok kontrol dan kelompok eksperimen menghasilkan $t = 1.681$ ($p < 0.05$, *one tailed*). Hipotesis penelitian kedua diterima. Terdapat perbedaan hasil tes CFIT A ditinjau dari perbedaan penggunaan jenis lembar jawaban. Subjek yang mengguna-

kan lembar jawaban manual mendapatkan hasil tes CFIT A yang lebih tinggi daripada subjek yang menggunakan lembar jawaban komputer.

Independent sample t-test terhadap angka inteligensi hasil skala CFIT 3B pada kelompok kontrol dan kelompok eksperimen menghasilkan $t = 2.289$ ($p < 0.05$, *one tailed*). Hipotesis penelitian ketiga diterima. Terdapat perbedaan hasil tes CFIT B ditinjau dari perbedaan penggunaan jenis lembar jawaban. Subjek yang menggunakan lembar jawaban manual mendapatkan hasil tes CFIT B yang lebih tinggi daripada subjek yang menggunakan lembar jawaban komputer.

Independent sample t-test terhadap angka inteligensi hasil skala CFIT 3A dan B pada kelompok kontrol dan kelompok eksperimen menghasilkan $t = 1.819$ ($p < 0.05$, *one tailed*). Hipotesis penelitian keempat diterima. Terdapat perbedaan hasil tes CFIT A dan B ditinjau dari perbedaan penggunaan jenis lembar jawaban. Subjek yang menggunakan lembar jawaban manual mendapatkan hasil tes CFIT A dan B yang lebih tinggi daripada subjek yang menggunakan lembar jawaban komputer.

Hasil analisis data menunjukkan bahwa semua hipotesis yang diajukan dalam penelitian ini teruji. Hal ini menunjukkan bahwa perbedaan penggunaan jenis lembar jawaban mempengaruhi waktu pengisian lembar jawaban dan hasil tes inteligensi. Waktu yang dibutuhkan untuk mengisi lembar jawaban manual lebih cepat daripada waktu yang diperlukan untuk mengisi lembar jawaban komputer. Hasil tes inteligensi dari subjek yang mengerjakan skala CFIT dengan menggunakan lembar jawaban manual juga lebih tinggi secara signifikan, dibandingkan subjek yang mengerjakan skala CFIT dengan menggunakan lembar jawaban komputer. Hal ini karena penggunaan

lembar jawaban komputer membuat waktu pengetesan tidak dapat dipergunakan oleh subjek secara penuh untuk berpikir dan menyelesaikan soal. Sebagian waktu pengetesan justru harus dipergunakan oleh subjek untuk menghitamkan bulatan alternatif jawaban yang dipilihnya. Hal ini perlu ditindak lanjuti dengan melakukan adaptasi waktu administrasi pelaksanaan tes inteligensi yang mempergunakan lembar jawaban komputer. Penelitian lebih lanjut diperlukan untuk melakukan adaptasi tersebut.

Jika praktisi ingin tetap menggunakan lembar jawaban komputer tanpa harus melakukan penelitian adaptasi waktu, peneliti menyarankan untuk menggunakan lembar jawaban komputer yang dapat dikerjakan tanpa harus menghitamkan bulatan alternatif jawaban. Hal ini karena sekarang sudah tersedia *scanner* yang mampu membaca coretan atau karakter huruf pada lembar jawaban komputer secara langsung.

Hal lain yang dapat dilakukan adalah dengan melakukan tes dengan menggunakan model *computer-based* seperti yang dilakukan oleh Kumar dan Helgeson (1995). Model ini dapat diwujudkan dengan cara mentransfer alat tes dan soal-soal di dalamnya ke dalam bentuk *software*. *Testi* dapat mengerjakan tes secara langsung di komputer dengan cara mengklik alternatif jawaban yang tersedia.

Peneliti menyarankan kepada para peneliti selanjutnya untuk melakukan penelitian terhadap tes psikologis lain yang juga diadministrasikan dengan menggunakan lembar jawaban komputer, khususnya yang mengharuskan *testi* untuk menghitamkan bulatan alternatif jawaban yang dipilihnya. Tes psikologis tersebut, di sisi lain, sebenarnya pada prosedur administrasi standar tes cukup dikerjakan dengan cara mencoret alternatif jawaban.

Penelitian-penelitian semacam ini akan membuka luas adaptasi administrasi standar tes psikologis. Adaptasi tersebut akan membuat *testi* tidak dirugikan, dan alat tes terjaga validitasnya.

Kesimpulan

Terbuktinya seluruh hipotesis yang diajukan dalam penelitian ini menunjukkan pengaruh perbedaan penggunaan jenis lembar jawaban terhadap angka inteligensi hasil skala CFIT. Oleh karena itu, para praktisi yang ingin menggunakan lembar jawaban komputer hendaknya melakukan adaptasi waktu administrasi tes dengan cara melakukan penelitian terlebih dahulu, atau menggunakan *scanner* yang sesuai. Pengabaian hal ini membuat hasil tes tidak dapat dipercaya, karena akan menghasilkan angka inteligensi yang *under estimate*.

Kepustakaan

- Anastasi, A., & Urbina, U. (2007). *Tes psikologi* (terjemahan). Jakarta: PT. Indeks.
- Austin, E.J., Hofer, S.M., Deary, I.J., & Eber, H.W. (2000). Interactions between intelligence and personality: results from two large samples. *Personality and Individual Differences*, 29(3), 405–427.
- Azwar, S. (2009). Reliabilitas dan validitas. Yogyakarta: Pustaka Pelajar.
- Azwar, S. (2010). *Pengantar psikologi inteligensi*. Yogyakarta: Pustaka Pelajar.
- Baron, R.A. (1996). *Essentials of psychology*. Needham Heights, MA: Allyn & Bacon.
- Cattell, R. B. (1949). *Culture fair intelligence test, scale 1*. Champaign, Illinois: Institute of Personality and Ability.
- Cattell, R.B. (1961). *Measuring intelligence*

- with the culture fair test*. Cambridge, MA: Hogrefe Publishing.
- Cattell, R.B. (1971). *Abilities: Their structure growth and action*. Boston: Houghton Mifflin.
- Cattell, R.B., & Cattell, A.K.S. (1973). *Culture fair intelligence tests (CFIT)*. San Diego, CA: Educational and Industrial Testing Service (EdITS).
- Colom, R., & Garcia-Lopez, O. (2002). Sex differences in fluid intelligence among high school graduates. *Personality and Individual Differences, 32*(3), 445–451.
- Colom, R., & Garcia-Lopez, O. (2003). Secular gains in fluid intelligence: Evidence from the culture-fair intelligence test. *Journal of Biosocial Science, 35*, 33–39.
- Fisteus, J.A., Pardo, A., & Garcia, N.F. (2012). Grading multiple choice exams with low-cost and portable computer-vision techniques. *Journal of Science, Education, and Technology*, DOI 10.1007/s10956-012-9414-8.
- Furlow, F.B., Armijo-Prewitt, T., Gangestad, S.W., & Thornhill, R. (1997). Fluctuating Asymmetry and Psychometric Intelligence. *Proceedings of Biological Sciences, 264*(1383), 823–829.
- Goldstein, G., & Hersen, M. (2000). *Handbook of psychological assessment*. Oxford: Pergamon Press.
- Gupta, S. (1991). Effects of time of day and personality on intelligence test scores. *Personality and Individual Differences, 12*(11), 1227–1231.
- Kumar, D.D., & Helgeson, S.L. (1995). Trends in computer applications in science assessment. *Journal of Science, Education, and Technology, 4*(1), 29–36.
- Myers, A. (1987). *Experimental psychology*. California: Brooks/Cole Publishing Company.
- Rammsayer, T.H., & Brandler, S. (2002). On the relationship between general fluid intelligence and psychophysical indicators of temporal resolution in the brain. *Journal of Research in Personality, 36*(5), 507–530.
- SMR Indonesia. (2012). *Deskripsi singkat SMR*. Diunduh dari: SMR Indonesia http://smrindonesia.com/index.php?option=com_content&task=viewed&id=45&Itemid=146 tanggal 24 Februari 2012.