

Pengaruh Adaptasi Waktu Administrasi yang disebabkan Penggunaan Lembar Jawaban Komputer terhadap Hasil CFIT 3 A dan 3 B

*Ridwan Saptoto*¹

Fakultas Psikologi Universitas Gadjah Mada

Abstract. Previous research showed difference in intelligence level caused by difference in answering method. This research aimed to adapt test time administered using computer answer sheet. Adaptation time was conducted by comparing time needed to give answer by making cross line and blackening answer dots using 2B computer pencil. Experiment method with two independent group design had chosen. Data analysis showed no difference in raw score between control group and experiment group in CFIT 3A ($t = 1.045, p > 0.05$), CFIT 3B ($t = 0.368, p > 0.05$), and CFIT 3A and 3B ($t = 0.791, p > 0.05$). Data analysis also found no difference in intelligence level between both groups in CFIT 3A ($t = 1.063, p > 0.05$), CFIT 3B ($t = 0.470, p > 0.05$), and CFIT 3A and 3B ($t = 0.771, p > 0.05$). Adaptation administration time did not change intelligence test results.

Keywords: CFIT; computer answer sheet; intelligence test

Abstrak. Hasil penelitian sebelumnya menunjukkan adanya perbedaan tingkat inteligensi yang disebabkan oleh perbedaan metode memberikan jawaban. Penelitian ini bertujuan melakukan adaptasi terhadap tes yang menggunakan lembar jawaban komputer. Adaptasi waktu dilakukan dengan membandingkan waktu yang dibutuhkan untuk memberikan jawaban dengan cara menyilang dan menghitamkan alternatif jawaban menggunakan pensil 2B. Metode eksperimen dengan desain dua kelompok independen dipilih dalam penelitian ini. Analisis data menunjukkan tidak terdapat perbedaan skor mentah antara kelompok kontrol dan kelompok eksperimen pada CFIT 3A ($t = 1,045, p > 0,05$), CFIT 3B ($t = 0,368, p > 0,05$), serta CFIT 3A dan 3B ($t = 0,791, p > 0,05$). Analisis data juga menemukan tidak terdapat perbedaan tingkat inteligensi antara kedua kelompok pada CFIT 3A ($t = 1,063, p > 0,05$), CFIT 3B ($t = 0,470, p > 0,05$), serta CFIT 3A dan 3B ($t = 0,771, p > 0,05$). Adaptasi waktu administrasi tidak mengubah hasil tes inteligensi.

Kata kunci: CFIT; lembar jawaban komputer; tes inteligensi

Ilmu psikologi sebagai ilmu pengetahuan ilmiah harus dapat diuji dan dibuktikan kebenarannya. Salah satu cara untuk mengujinya yaitu dengan menggunakan sebuah alat ukur. Hal ini membuat

perkembangan ilmu psikologi pada akhirnya akan selalu diikuti dengan perkembangan alat ukurnya.

Jika pada zaman dahulu pengukuran dalam ilmu psikologi banyak dilakukan dengan menggunakan *paper and pencil test*, maka pada saat ini pengukuran dalam

¹ Korespondensi mengenai artikel ini dapat melalui:
ridwan_psychology@ugm.ac.id

ilmu psikologi sudah menggunakan komputer. Kumar dan Helgeson (1995) yang melakukan penelitian mengenai penggunaan komputer dalam pengetesan menemukan bahwa metode ini mampu memperbaiki efektivitas dan efisiensi dalam pencatatan, penyimpanan, dan analisis data.

Salah satu bentuk pemanfaatan komputer dalam pengetesan yaitu dengan digunakannya Lembar Jawab Komputer (selanjutnya disingkat LJK). Pada metode ini peserta tes mengerjakan tes atau memberikan jawaban pada LJK dengan cara menghitamkan bulatan dari alternatif jawaban yang dipilihnya. LJK kemudian di-scan dengan menggunakan sebuah alat khusus yang terhubung dengan komputer. Komputer yang sudah ter-install program beserta kunci jawaban tes akan langsung memeriksa jawaban peserta tes, sehingga hasil tes dapat diketahui secara cepat. LJK biasanya digunakan untuk tes-tes yang diikuti oleh banyak peserta. Jika pemeriksaan tersebut dilakukan secara manual, maka akan membutuhkan waktu yang lama. Program-program komputer untuk beberapa jenis tes psikologis yang dikerjakan dengan menggunakan LJK bahkan telah tersedia dan diperjualbelikan oleh sebuah perusahaan (SMR Indonesia, 2012).

Tes psikologis merupakan alat ukur objektif yang telah distandardisasikan atau dibakukan cara dan ketentuan pengukurannya. Hasil tes psikologis dapat menunjukkan berbagai aspek psikologis dari mereka yang mengerjakannya (Anastasi & Urbina, 2007). Berbagai tes psikologis tersedia untuk mengungkap aspek-aspek psikologis seseorang, antara lain yaitu aspek inteligensi, bakat, minat, dan kepribadian.

Azwar (2010) mendefinisikan tes inteligensi sebagai tes yang mengungkap kemampuan mental individu. Ahli-ahli

lain memberikan definisi yang beragam tentang inteligensi, sehingga pengembangan tes inteligensi pun menjadi beragam. Beberapa ahli tersebut kemudian ada yang mengembangkan tes inteligensi berdasarkan definisi inteligensi yang disusunnya sendiri. Salah satu tes inteligensi yang berkembang dan saat ini masih digunakan adalah *Culture Fair Intelligence Test* (selanjutnya disingkat CFIT). Tes ini dikembangkan oleh Cattell berdasarkan pandangannya mengenai inteligensi. Alat ini mengukur *general intelligence* (Cattell, 1949).

Cattell (1971) menyebut inteligensi sebagai *general intelligence*. Inteligensi terdiri dari dua faktor, yaitu *fluid intelligence* dan *crystalized intelligence*. *Fluid intelligence* adalah kemampuan seseorang untuk menggunakan logikanya dalam menyelesaikan permasalahan dan mengumpulkan informasi yang dibutuhkan untuk menyelesaikan permasalahan tersebut. Permasalahan tersebut tidak hanya berupa permasalahan yang sudah pernah dihadapi sebelumnya, namun juga permasalahan-permasalahan baru yang menuntut penyelesaian dengan cara berpikir logis.

Cattell (1971) selanjutnya menyebutkan bahwa *crystalized intelligence* merupakan kemampuan seseorang untuk memanfaatkan segenap keterampilan, pengetahuan, dan pengalaman yang telah dimilikinya secara optimal. Kemampuan ini terkait dengan bagaimana seseorang mampu memanfaatkan ingatan jangka panjang yang dimilikinya. Kemampuan ini juga akan meningkat dengan semakin bertambahnya usia seseorang.

Cattell dan Cattell (1973) menyebutkan bahwa CFIT terdiri dari tiga jenis tes atau skala. Masing-masing skala diberi nama dengan angka, sehingga nama-nama skala tersebut adalah skala 1, skala 2, dan

skala 3. Skala 1 dapat dimanfaatkan untuk mengungkap kecerdasan anak-anak yang berumur empat sampai dengan delapan tahun, atau orang dengan usia yang lebih tua namun mengalami kesulitan belajar. Skala 2 dapat dimanfaatkan untuk mengungkap kecerdasan anak-anak yang berusia delapan sampai dengan empat belas tahun, atau orang dewasa yang memiliki kecerdasan normal. Terakhir yaitu skala 3. Skala ini dapat dimanfaatkan untuk mengungkap kecerdasan orang yang berusia empat belas tahun ke atas, atau orang dewasa yang memiliki taraf kecerdasan superior. Skala CFIT 2 dan 3 memiliki bentuk paralel, yaitu A dan B. Hal ini berarti bahwa skala CFIT 2 terdiri dari 2A dan 2B, sedangkan skala CFIT 3 terdiri dari 3A dan 3B.

Cattell dan Cattell (1973) selanjutnya mengungkapkan bahwa skala CFIT 3 terdiri dari empat subtes. Keempat subtes tersebut yaitu *series*, *classification*, *matrices*, dan *conditions* atau *topology*. Setiap subtes hanya boleh dikerjakan dalam waktu yang telah distandardisasikan. Peserta tes atau *teste* mulai mengerjakan tes setelah mendapatkan tanda dari *tester*. *Teste* kemudian harus meletakkan alat tulisnya ketika *tester* memerintahkan untuk berhenti karena waktu yang ditentukan telah habis. CFIT 3 dikerjakan dengan cara memberikan satu garis diagonal atau satu garis coret pada pilihan jawaban yang dianggap benar. Tanda garis diagonal tersebut diberikan pada lembar jawaban, bukan pada buku soal CFIT. Tes CFIT pada saat ini sudah dapat dikerjakan dengan menggunakan lembar jawaban komputer (SMR Indonesia, 2012). Penggunaan jenis lembar jawaban tersebut membuat *teste* harus menghitamkan bulatan pada alternatif jawaban yang dipilihnya.

Hal ini tentunya membutuhkan tambahan waktu dalam pemilihan alternatif jawaban, dan dengan sendirinya akan mengurangi waktu *teste* untuk memikirkan alternatif jawaban yang benar. *Tester* sendiri pada umumnya tidak memberikan waktu tambahan atau memberikan waktu khusus bagi *teste* untuk menghitamkan bulatan alternatif jawaban yang dipilihnya. *Tester* tetap berpedoman dan menggunakan standar waktu yang sama dengan cara pemberian jawaban yang dicoret. Kondisi di atas juga terjadi dalam proses pemberian tes psikologis yang mempergunakan lembar jawaban komputer. Hal ini berpotensi memengaruhi hasil pengetesan.

Saptoto (2012) dalam penelitiannya menemukan bahwa terdapat perbedaan waktu yang signifikan antara kelompok subjek yang mengerjakan CFIT dengan cara mencoret alternatif jawaban dan kelompok subjek yang mengerjakan CFIT dengan cara menghitamkan bulatan alternatif jawaban. Kelompok subjek pertama membutuhkan waktu yang lebih cepat daripada kelompok subjek kedua.

Hasil penelitian yang dilakukan oleh Saptoto (2012) juga menemukan bahwa terdapat perbedaan hasil CFIT yang signifikan antara kelompok subjek yang mengerjakan CFIT dengan cara mencoret alternatif jawaban dan kelompok subjek yang mengerjakan CFIT dengan cara menghitamkan bulatan alternatif jawaban. Kedua kelompok subjek tersebut mengerjakan CFIT dengan menggunakan waktu yang sama, yaitu waktu standar administrasi CFIT yang berlaku. Kelompok subjek pertama secara konsisten mendapatkan hasil CFIT yang lebih tinggi daripada kelompok subjek kedua. Kelompok subjek pertama selalu mendapatkan hasil CFIT yang lebih tinggi dibandingkan kelompok subjek kedua pada pengetesan dengan

menggunakan CFIT 3A dan CFIT 3B. Bahkan jika hasil CFIT 3A dan 3B digabungkan, kelompok subjek pertama tetap masih mendapatkan hasil CFIT yang lebih tinggi daripada kelompok subjek kedua. Hal ini menunjukkan adanya pengaruh jenis lembar jawaban terhadap hasil CFIT. Oleh karena itu diperlukan adanya adaptasi waktu administrasi CFIT yang menggunakan lembar jawaban komputer.

International Test Commission (selanjutnya disingkat ITC) dibentuk pada tahun 1991 untuk memenuhi kebutuhan akan adaptasi tes. Komisi tersebut kemudian menyusun panduan dalam penerjemahan dan adaptasi tes supaya hasilnya dapat dipertanggungjawabkan. Panduan ini dipublikasikan pada tahun 1994 (Gregoire & Hambleton, 2009). Namun demikian, penggunaan internet dan komputer dalam penyelenggaraan tes telah mengubah proses penyajian, penilaian, dan pelaporan hasil tes. Hal ini memunculkan kebutuhan untuk mengubah tata cara dalam penyajian tes kepada para pengambil tes (Hambleton, Merenda & Spielberger, 2005). ITC pada tahun 2005 kemudian membentuk komite baru untuk memperbaiki tata cara penerjemahan dan adaptasi sebuah tes (Gregorie & Hambleton, 2009). Komite ini kemudian menghasilkan tata cara baru dalam penerjemahan dan adaptasi tes pada tahun 2016.

Panduan ITC (2017) menyebutkan bahwa ada dua hal yang perlu dilakukan oleh peneliti pada saat akan menerjemahkan dan mengadaptasi tes. Pertama, peneliti harus menyiapkan semua material dan instruksi secara seksama untuk meminimalisasi masalah-masalah terkait budaya dan bahasa yang disebabkan oleh prosedur administrasi dan cara pemberian respon, sehingga memengaruhi validitas atas kesimpulan yang ditarik dari skor tes.

ITC selanjutnya menjelaskan bahwa masalah-masalah lain dalam administrasi tes yang dapat memengaruhi hasil tes juga perlu diperhatikan, termasuk di dalamnya adalah mekanisme dalam memberikan jawaban (misalnya bentuk lembar jawaban), dan waktu yang diperlukan untuk menyelesaikan tes. ITC menekankan pentingnya waktu dalam administrasi tes, karena kesalahan yang jamak terjadi yaitu pengambil tes tidak memiliki waktu yang cukup untuk menyelesaikan sebuah tes.

Hal kedua yang harus dilakukan dalam proses penerjemahan dan adaptasi tes adalah menjelaskan kondisi tes yang seharusnya diikuti secara cermat, termasuk batasan waktu dalam pengerjaan tes (ITC, 2017). Tujuan dari aturan ini adalah untuk menjamin pengambil tes dapat menunjukkan pengetahuan, kemampuan, dan keterampilan yang dimilikinya secara sesungguhnya

Berdasarkan uraian-uraian di atas, maka dapat disimpulkan bahwa adaptasi administrasi CFIT perlu dilakukan agar validitas hasil tes dapat dipercaya. Peneliti selanjutnya mengajukan enam hipotesis di dalam penelitian ini, yaitu: 1) Tidak terdapat perbedaan hasil *raw score* CFIT 3A antara kelompok subjek yang menggunakan waktu standar administrasi pada saat mengerjakan CFIT 3A, dengan kelompok subjek yang menggunakan adaptasi waktu standar administrasi pada saat mengerjakan CFIT 3A. 2) Tidak terdapat perbedaan hasil *raw score* CFIT 3B antara kelompok subjek yang menggunakan waktu standar administrasi pada saat mengerjakan CFIT 3B, dengan kelompok subjek yang menggunakan adaptasi waktu standar administrasi pada saat mengerjakan CFIT 3B. 3) Tidak terdapat perbedaan hasil *raw score* CFIT 3A + 3B antara kelompok subjek yang menggunakan waktu

standar administrasi pada saat mengerjakan CFIT 3A + 3B, dengan kelompok subjek yang menggunakan adaptasi waktu standar administrasi pada saat mengerjakan CFIT 3A + 3B. 4) Tidak terdapat perbedaan inteligensi dari CFIT 3A antara kelompok subjek yang menggunakan waktu standar administrasi pada saat mengerjakan CFIT 3A, dengan kelompok subjek yang menggunakan adaptasi waktu standar administrasi pada saat mengerjakan CFIT 3A. 5) Tidak terdapat perbedaan inteligensi dari CFIT 3B antara kelompok subjek yang menggunakan waktu standar administrasi pada saat mengerjakan CFIT 3B, dengan kelompok subjek yang menggunakan adaptasi waktu standar administrasi pada saat mengerjakan CFIT 3B. 6) Tidak terdapat perbedaan inteligensi dari CFIT 3A + 3B antara kelompok subjek yang menggunakan waktu standar administrasi pada saat mengerjakan CFIT 3A + 3B, dengan kelompok subjek yang menggunakan adaptasi waktu standar administrasi pada saat mengerjakan CFIT 3A + 3B.

Metode

Subjek penelitian

Subjek penelitian adalah para mahasiswa sebuah Perguruan Tinggi X yang mendaftarkan diri dan bersedia mengikuti prosedur eksperimen yang telah disusun. Peneliti mendapatkan 130 orang mahasiswa yang bersedia menjadi subjek penelitian dari berbagai macam daerah di Indonesia. Semua subjek penelitian dimasukkan ke dalam kelompok kontrol atau kelompok eksperimen dengan menggunakan metode *random assignment*. Myers dan Hansen (2002) mengatakan bahwa *random assignment* memungkinkan setiap subjek memiliki kesempatan yang

sama untuk ditempatkan ke dalam berbagai kondisi perlakuan.

Hasil *random assignment* menempatkan 65 orang subjek di kelompok kontrol, dan sisanya menjadi subjek di kelompok eksperimen. Namun demikian, pada saat pengambilan data terdapat dua orang di kelompok eksperimen dan tiga orang di kelompok kontrol yang tidak hadir. Hal ini membuat jumlah subjek di kelompok eksperimen menjadi 63 orang, sedangkan subjek di kelompok kontrol berjumlah 62 orang.

Para subjek yang mendaftar memiliki rentang usia antara 20 sampai dengan 22 tahun, tersebar dari mereka yang menduduki bangku kuliah di semester 3, semester 5, dan semester 7. Jumlah subjek laki-laki di kelompok eksperimen adalah 18 orang, dan sisanya sebanyak 45 orang adalah perempuan. Selanjutnya jumlah subjek laki-laki di kelompok kontrol adalah 16 orang, dan sisanya sebanyak 46 orang adalah perempuan

Instrumen penelitian

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

Data penelitian kedua berupa *raw score* dan inteligensi. Data inteligensi diambil dengan menggunakan CFIT 3A dan 3B yang disusun oleh Cattell dan Cattell (1973). Tes ini mengukur *fluid intelligence*, yaitu kemampuan seseorang untuk menggunakan logikanya dalam

menyelesaikan permasalahan dan mengumpulkan informasi yang dibutuhkan untuk menyelesaikan permasalahan tersebut. CFIT 3A dan 3B terdiri dari empat subtes yang memiliki waktu pengerjaannya sendiri-sendiri, yaitu subtes *series*, *classification*, *matrices*, dan *conditions*. Subtes *series* terdiri dari 13 aitem dan waktu pengerjaannya adalah 3 menit. Selanjutnya subtes *classification* tersusun dari 14 aitem, dan waktu pengerjaannya selama 4 menit. Subtes ketiga, yaitu *matrices*, terdiri dari 13 aitem, dan perlu dikerjakan selama 3 menit. Terakhir yaitu subtes *conditions* yang terdiri dari 10 aitem, dan memiliki waktu pengerjaan sebanyak 2,5 menit.

Hasil korelasi CFIT dengan berbagai tes inteligensi menunjukkan koefisien korelasi $r=0,5$ sampai dengan $r=0,7$ (Goldstein & Hersen, 2000). Koefisien korelasi tersebut membuktikan bahwa CFIT valid untuk mengukur inteligensi. Tes yang valid berarti bahwa tes tersebut mampu melaksanakan fungsi pengukurannya sesuai tujuan disusunnya atau diberikannya tes tersebut kepada subjek (Azwar, 2009).

Hasil penelitian yang dilakukan oleh Goldstein dan Hersen (2000) selanjutnya juga menemukan bahwa CFIT skala 3 memiliki koefisien konsistensi internal sebesar 0,85, sedangkan reliabilitas tes ulang CFIT skala 3 menghasilkan koefisien korelasi $r=0,82$. Hasil penelitian tersebut menunjukkan bahwa CFIT skala 3 ternyata memiliki reliabilitas yang cukup tinggi untuk mengukur inteligensi. Reliabilitas merupakan keterandalan atau keajegan sebuah tes. Tes yang memiliki reliabilitas tinggi atau reliabel menandakan bahwa tes tersebut mampu menunjukkan hasil penggetesan yang relatif sama dalam beberapa pengukuran terhadap subjek

atau kelompok subjek yang sama (Azwar, 2009).

Prosedur

Peneliti pada awalnya membuka kesempatan untuk menjadi subjek penelitian kepada mahasiswa di Fakultas Psikologi sebuah Perguruan Tinggi X. Informasi mengenai hal ini disampaikan melalui poster maupun sistem informasi akademik. Peneliti kemudian mencatat ketersediaan waktu yang mereka miliki untuk pengambilan data, dan sekaligus menyusun jadwal pengambilan data yang dapat diakomodasi oleh para subjek penelitian.

Desain penelitian yang dipergunakan dalam penelitian ini adalah desain eksperimen. Penelitian ini sendiri terdiri dari dua tahap. Penelitian tahap pertama bertujuan untuk mengidentifikasi perbedaan waktu yang diperlukan untuk memberikan jawaban dengan cara mencoret dan menghitamkan bulatan pada alternatif jawaban yang sesuai. Perbedaan waktu yang ditemukan kemudian digunakan sebagai dasar untuk melakukan adaptasi waktu administrasi CFIT 3A dan 3B dengan menggunakan lembar jawaban komputer. Penelitian tahap kedua bertujuan untuk membuktikan bahwa adaptasi waktu administrasi yang telah dilakukan dalam pengerjaan tes dengan menggunakan LJK tidak mengubah hasil tes. Hal ini karena pemberian jawaban dengan menggunakan LJK membutuhkan waktu yang lebih banyak, sedangkan pemberian jawaban pada lembar jawaban manual sesuai manual asli CFIT hanya membutuhkan waktu yang sangat singkat.

Seluruh subjek penelitian mengikuti eksperimen tahap pertama yang terdiri dari dua aktivitas. Para subjek pada aktivitas pertama diminta untuk memberikan jawaban pada lembar jawaban

manual dengan cara memberikan satu garis diagonal atau satu garis coret pada pilihan jawaban. Setiap nomor pada lembar jawaban manual hanya perlu dicoret sekali pada sebuah alternatif pilihan jawaban yang sudah ditentukan sebelumnya. Penentuan alternatif jawaban tersebut dilakukan secara random sebelum penelitian dimulai.

Peneliti berikutnya memberikan LJK setelah semua subjek menyelesaikan tugasnya yang pertama. Para subjek lalu diberi tugas untuk menghitamkan bulatan pada alternatif jawaban yang telah ditentukan pada LJK. Setiap nomor pada LJK hanya perlu dihitamkan bulatannya satu kali. Alternatif jawaban yang dihitamkan bulatannya sama dengan alternatif jawaban pada tugas sebelumnya. Eksperimen tahap pertama selesai sampai di sini.

Semua subjek pada eksperimen pertama maupun kedua sama sekali tidak mengerjakan CFIT 3A maupun 3B. Mereka hanya mencoret lembar jawaban manual pada aktivitas pertama, selanjutnya menghitamkan bulatan alternatif jawaban pada aktivitas kedua. Waktu yang diperlukan oleh setiap subjek untuk menyelesaikan aktivitas pertama maupun kedua dicatat dan dihitung, demikian juga dengan selisih waktu di antara keduanya.

Eksperimen berlanjut ke tahap kedua. Para subjek dikelompokkan secara random ke dalam kelompok kontrol atau kelompok eksperimen. Kelompok kontrol pada eksperimen tahap kedua sesi pertama diberi tugas untuk mengerjakan CFIT 3A dengan cara memberikan coretan sekali pada alternatif jawaban yang dipilihnya di lembar jawaban. Waktu yang digunakan untuk mengerjakan tes ini adalah waktu standar administrasi CFIT 3A. Mereka selanjutnya pada tahap kedua sesi kedua kembali diberi tugas untuk

mengerjakan CFIT 3B dengan cara memberikan coretan sekali pada alternatif jawaban yang dipilihnya di lembar jawaban. Tes juga dilaksanakan dengan menggunakan waktu standar administrasi.

Kelompok eksperimen pada eksperimen tahap kedua sesi pertama diberi tugas untuk mengerjakan CFIT 3A dengan cara menghitamkan bulatan pada alternatif jawaban yang dipilihnya di LJK. Waktu yang digunakan untuk mengerjakan tes ini bukanlah waktu standar administrasi, namun adaptasi waktu standar administrasi. Mereka selanjutnya pada tahap kedua sesi kedua kembali diberi tugas untuk mengerjakan CFIT 3B dengan cara menghitamkan bulatan pada alternatif jawaban yang dipilihnya di LJK. Tes juga dilaksanakan dengan menggunakan adaptasi waktu standar administrasi.

Jawaban semua subjek pada saat mereka mengerjakan CFIT 3A dan 3B, baik yang berasal dari kelompok kontrol maupun kelompok eksperimen, kemudian dikoreksi. Jawaban-jawaban benar dari setiap soal kemudian dijumlahkan, sehingga menghasilkan *raw score*. *Raw score* ini selanjutnya dikonversikan menjadi angka inteligensi dengan menggunakan norma CFIT 3A dan 3B.

Pengambilan data penelitian tahap pertama dilakukan terhadap 125 mahasiswa yang mendaftar sebagai subjek penelitian dan hadir pada saat pengambilan data. Waktu yang dibutuhkan oleh setiap subjek untuk memberikan jawaban dengan cara menyilang atau menghitamkan bulatan alternatif jawaban pada lembar jawaban yang sesuai diukur dengan menggunakan *stopwatch*.

Penelitian tahap pertama menggunakan desain eksperimen yang disebut sebagai *simple design* (Myers & Hansen, 2002). 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. Jarak waktu antara pengukuran pertama dan kedua adalah 15 menit. Hal ini dilakukan untuk mencegah terjadinya *fatigue effects*, yaitu penurunan performa subjek karena mengalami kelelahan setelah mengikuti beberapa eksperimen secara terus menerus (Myers & Hansen, 2002). Desain penelitian tahap pertama yang digunakan oleh peneliti dapat dilihat pada Gambar 1.

KE	X1O1	X2O2
----	------	------

Gambar 1. Desain Penelitian Tahap Pertama

Keterangan:

KE : kelompok eksperimen

X_{1O1} : subjek menyilang alternatif jawaban, dan peneliti menghitung waktu yang diperlukan untuk memilih dan menyilang alternatif jawaban sesuai nomor soal pada lembar jawaban

X_{2O2} : subjek menghitamkan bulatan alternatif jawaban, dan peneliti menghitung waktu yang diperlukan untuk memilih dan menghitamkan bulatan alternatif jawaban sesuai nomor soal pada lembar jawaban

Pengambilan data penelitian tahap kedua dilakukan satu minggu setelah seluruh subjek menyelesaikan eksperimen tahap pertama. Hal ini karena eksperimen tahap pertama membutuhkan waktu yang lebih intensif. Setiap subjek harus didampingi secara individual pada saat mencoret lembar jawaban manual maupun menghitamkan bulatan alternatif jawaban, sebab masing-masing aktivitas waktunya harus dicatat secara teliti dengan menggunakan *stopwatch*.

Penelitian pada tahap eksperimen kedua menggunakan *two independent group design* (Myers & Hansen, 2002). Peneliti pada tahap ini menggunakan dua jenis kelompok, yaitu kelompok kontrol dan kelompok eksperimen. Subjek kelompok kontrol berbeda dengan subjek kelompok eksperimen. Masing-masing kelompok mendapatkan jenis perlakuan yang berbeda.

Penempatan subjek ke dalam kelompok kontrol atau kelompok eksperimen dilakukan dengan menggunakan *random assignment* atau acak. Menurut Myers dan Hansen (2002) *random assignment* akan membuat setiap subjek memiliki kesempatan yang sama untuk ditempatkan ke dalam kelompok kontrol ataupun kelompok eksperimen. Hal ini akan membuat kondisi awal kedua kelompok berada pada tingkatan yang setara, sehingga keduanya juga dapat diperbandingkan secara setara. Jarak waktu antara pengukuran pertama dan kedua adalah 15 menit. Hal ini dilakukan untuk mencegah terjadinya *fatigue effects*. Subjek pada kelompok kontrol dan eksperimen mengerjakan tes di ruang yang sama, namun waktunya berbeda. Desain penelitian tahap kedua yang dipergunakan oleh peneliti disajikan dalam Gambar 2.

	KK	X1O1	X2O2
R			
	KE	X3O3	X4O4

Gambar 2. Desain Penelitian Tahap Kedua

Keterangan:

R : random

KK : kelompok kontrol

KE : kelompok eksperimen

X_{1O1} : subjek mengerjakan CFIT 3A dengan mencoret alternatif jawaban dan peng-

- ambilan data inteligensi oleh peneliti, dengan menggunakan waktu standar administrasi
- X₂O₂ : subjek mengerjakan CFIT 3B dengan mencoret alternatif jawaban dan pengambilan data inteligensi oleh peneliti, dengan menggunakan waktu standar administrasi
- X₃O₃ : subjek mengerjakan CFIT 3A dengan menghitamkan bulatan alternatif jawaban dan pengambilan data inteligensi oleh peneliti, dengan menggunakan adaptasi waktu standar administrasi
- X₄O₄ : subjek mengerjakan CFIT 3B dengan menghitamkan bulatan alternatif jawaban dan pengambilan data inteligensi oleh peneliti, dengan menggunakan adaptasi waktu standar administrasi

Hipotesis-hipotesis penelitian yang diajukan dalam penelitian ini akan diuji dengan menggunakan metode statistik. Metode analisis data yang digunakan untuk menguji keenam hipotesis penelitian adalah *independent sample t-test*. Pengujian statistik 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

Pengambilan data tahap pertama menunjukkan bahwa *mean* waktu yang

dibutuhkan subjek untuk memberikan jawaban dengan cara menyilang alternatif jawaban pada setiap nomor adalah 0,926 detik. Selanjutnya *mean* waktu yang dibutuhkan subjek untuk memberikan jawaban dengan cara menghitamkan bulatan alternatif jawaban pada setiap soal adalah 4,370 detik. Oleh karena itu, tambahan waktu yang diperlukan untuk setiap soal yang dikerjakan dengan cara menghitamkan bulatan alternatif jawaban = 4,370 detik – 0,926 detik = 3,444 detik.

Tambahan waktu tersebut kemudian dikalikan dengan jumlah soal pada setiap subtes CFIT. Hasil pengkalian yang berupa total tambahan waktu kemudian ditambahkan pada waktu standar administrasi CFIT pada setiap subtes. Berdasarkan perhitungan tersebut, maka didapatkan adaptasi waktu standar administrasi CFIT yang disajikan pada Tabel 1.

Statistik deskriptif menunjukkan bahwa perbedaan rerata *raw score* antara kelompok kontrol dan kelompok eksperimen pada CFIT 3A, 3B, dan skor total CFIT 3A ditambah CFIT 3B tidak melebihi 1.5 nilai. Sedangkan perbedaan rerata *scaled score* antara kelompok kontrol dan kelompok eksperimen pada CFIT 3A, 3B, dan skor total CFIT 3A ditambah CFIT 3B tidak melebihi 3 nilai. Data-data mean dan selisihnya secara lebih terperinci dapat dilihat pada Tabel 2 berikut ini.

Tabel 1
Adaptasi waktu standar administrasi CFIT

Keterangan	Jumlah Item	Jumlah Jawaban	Waktu Standar	Adaptasi Waktu
Subtes 1	13	13	3'	3'45"
Subtes 2	14	28	4'	5'36"
Subtes 3	13	13	3'	3'45"
Subtes 4	10	10	2'30"	3'04"
Jumlah			12'30"	14'40"

Tabel 2

Mean dan perbedaan mean antara kelompok eksperimen dan kelompok kontrol

Jenis Data	Kelompok	Mean	Perbedaan Mean
Raw Score CFIT 3A	Eksperimen	26,9	0,82
	Kontrol	26,08	
Scaled Score CFIT 3A	Eksperimen	114,85	2,54
	Kontrol	112,31	
Raw Score CFIT 3A	Eksperimen	30,92	0,3
	Kontrol	30,62	
Scaled Score CFIT 3A	Eksperimen	114,12	1,15
	Kontrol	112,97	
Raw Score CFIT 3A + CFIT 3B	Eksperimen	57,82	1,12
	Kontrol	56,7	
Scaled Score CFIT 3A + CFIT 3B	Eksperimen	115,63	1,79
	Kontrol	113,84	

Independent sample t-test terhadap *raw score* hasil skala CFIT 3A pada kelompok kontrol dan kelompok eksperimen menghasilkan $t = 1,045$ ($p > 0,05$). Hipotesis penelitian pertama diterima. Tidak terdapat perbedaan hasil *raw score* CFIT 3A antara kelompok subjek yang menggunakan waktu standar administrasi pada saat mengerjakan CFIT 3A, dengan kelompok subjek yang menggunakan adaptasi waktu standar administrasi pada saat mengerjakan CFIT 3A.

Independent sample t-test terhadap *raw score* hasil skala CFIT 3B pada kelompok kontrol dan kelompok eksperimen menghasilkan $t = 0,368$ ($p > 0,05$). Hipotesis penelitian kedua diterima. Tidak terdapat perbedaan hasil *raw score* CFIT 3B antara kelompok subjek yang menggunakan waktu standar administrasi pada saat mengerjakan CFIT 3B, dengan kelompok subjek yang menggunakan adaptasi waktu standar administrasi pada saat mengerjakan CFIT 3B.

Independent sample t-test terhadap *raw score* hasil skala CFIT 3A dan 3B pada kelompok kontrol dan kelompok eksperimen menghasilkan $t = 0,791$ ($p > 0,05$). Hipotesis penelitian ketiga diterima. Tidak terdapat perbedaan hasil *raw score* CFIT 3A dan 3B antara kelompok subjek yang menggunakan waktu standar administrasi pada saat mengerjakan CFIT 3A dan 3B, dengan kelompok subjek yang menggunakan adaptasi waktu standar administrasi pada saat mengerjakan CFIT 3A dan 3B.

Independent sample t-test terhadap inteligensi hasil skala CFIT 3A pada kelompok kontrol dan kelompok eksperimen menghasilkan $t = 1,063$ ($p > 0,05$). Hipotesis penelitian keempat diterima. Tidak terdapat perbedaan inteligensi dari CFIT 3A antara kelompok subjek yang menggunakan waktu standar administrasi pada saat mengerjakan CFIT 3A, dengan kelompok subjek yang menggunakan adaptasi waktu standar administrasi pada saat mengerjakan CFIT 3A.

Independent sample t-test terhadap inteligensi hasil skala CFIT 3B pada kelompok kontrol dan kelompok eksperimen menghasilkan $t = 1,470$ ($p > 0,05$). Hipotesis penelitian kelima diterima. Tidak terdapat perbedaan inteligensi dari CFIT 3B antara kelompok subjek yang menggunakan waktu standar administrasi pada saat mengerjakan CFIT 3B, dengan kelompok subjek yang menggunakan adaptasi waktu standar administrasi pada saat mengerjakan CFIT 3B.

Independent sample t-test terhadap inteligensi hasil skala CFIT 3A dan 3B pada kelompok kontrol dan kelompok eksperimen menghasilkan $t = 0,771$ ($p > 0,05$). Hipotesis penelitian keenam diterima. Tidak terdapat perbedaan inteligensi dari CFIT 3A dan 3B antara kelompok subjek yang menggunakan waktu standar administrasi pada saat mengerjakan CFIT 3A dan 3B, dengan kelompok subjek yang menggunakan adaptasi waktu standar administrasi pada saat mengerjakan CFIT 3A dan 3B.

Diskusi

CFIT 3A dan 3B merupakan alat tes yang mengukur *general intelligence* (Cattell, 1949) pada orang yang memiliki usia empat belas tahun ke atas, atau orang dewasa yang memiliki taraf kecerdasan superior. Tes-tes tersebut sesungguhnya dikerjakan dengan cara memberikan satu garis diagonal atau satu garis coret pada pilihan jawaban yang dianggap benar di lembar jawaban. Namun demikian, pada saat ini baik CFIT 3A maupun 3B dapat dikerjakan dengan menggunakan lembar jawaban komputer (SMR Indonesia, 2012). Saptoto (2012) menemukan bahwa perbedaan administrasi CFIT ini menimbulkan perbedaan hasil CFIT yang signifikan antara kelompok subjek yang mengerjakan

CFIT dengan cara mencoret alternatif jawaban dan kelompok subjek yang mengerjakan CFIT dengan cara menghitamkan bulatan alternatif jawaban. Kelompok subjek pertama selalu mendapatkan hasil CFIT yang lebih tinggi dibandingkan kelompok subjek kedua, karena waktu yang seharusnya dipergunakan untuk berpikir pada kelompok subjek kedua juga harus dialokasikan untuk menghitamkan bulatan alternatif jawaban pada lembar jawaban komputer.

Peneliti lain yang meneliti perbedaan lembar jawaban terhadap hasil tes adalah Miller (1965). Miller meneliti pengaruh lembar jawaban IBM 85 dan IBM 1230 terhadap hasil tes. Hasilnya perbedaan lembar jawaban tersebut membuat hasil tes subjek-subjeknya berbeda pula.

Hasil-hasil penelitian di atas sesuai dengan hasil penelitian yang dilakukan oleh Boyle (1984). Boyle meneliti pengaruh empat jenis lembar jawaban komputer terhadap hasil tes bakat. Keempat lembar jawaban komputer tersebut berasal dari tiga merek berbeda, dan memiliki luasan ruang alternatif jawaban yang berbeda-beda pula. Hasil penelitiannya menemukan bahwa jenis lembar jawaban memengaruhi hasil tes bakat yang berjenis *speed test*, yaitu tes yang waktu pengerjaannya dibatasi. Kondisi ini sesuai dengan CFIT, karena waktu pengerjaan CFIT juga dibatasi dengan ketat.

Hasil-hasil penelitian di atas menunjukkan bahwa perbedaan jenis lembar jawaban dan cara menjawab sebuah tes ternyata dapat memengaruhi validitas dari alat ukur, sehingga, sesuai panduan dari ITC (2017), perlu dilakukan adaptasi tes. ITC menekankan pentingnya waktu dalam administrasi tes, karena pengambil tes tidak memiliki waktu yang cukup untuk menyelesaikan sebuah tes merupakan kesalahan yang umum terjadi.

Brothen (2012) selanjutnya melakukan penelitian terhadap batasan waktu yang dipergunakan pada tes. Hasil penelitiannya menunjukkan bahwa waktu yang dipergunakan untuk mengerjakan tes antara subjek yang hidup satu abad lalu dengan subjek pada masa kini tidak berbeda secara signifikan. Perbandingan tersebut dilakukan dengan mempergunakan jenis lembar jawaban yang sama. Hal ini menimbulkan implikasi bahwa alokasi waktu subjek ketika mengerjakan tes tidak boleh berkurang untuk mengerjakan hal-hal lainnya. Oleh karena itu, jenis soal yang sama namun dikerjakan dengan mempergunakan lembar jawaban yang berbeda tidak boleh memberikan hasil tes yang berbeda.

Berbagai penelitian telah dilakukan dalam topik bahasan terkait adaptasi waktu dan pengetesan yang dikerjakan dengan menggunakan komputer. Sie, Finkelman, Riley, dan Smits (2015) menggunakan *Fisher information per time unit* untuk meningkatkan efisiensi waktu pengerjaan tes, sedangkan Choe, Kern, dan Chang (2017) menggunakan informasi yang sama untuk memilih aitem pada tes-tes yang diadaptasi dengan menggunakan komputer. Namun demikian, penelitian yang dilakukan untuk mengadaptasi waktu tes sehingga memberikan hasil yang adil baik pada pengetesan dengan menggunakan lembar jawaban manual maupun lembar jawaban komputer belum pernah dilakukan. Penelitian ini mencoba untuk mengisi celah pengembangan tersebut, dan menggunakan metode yang tepat untuk mencapai tujuan penelitian. Metode eksperimen yang dilakukan dalam dua tahap dan menggunakan data yang diambil langsung dari kedua kelompok subjek eksperimen. Penggunaan metode dan tahapan eksperimen tersebut diyakini mampu mengidentifikasi perbedaan wak-

tu yang diperlukan untuk memberikan jawaban dengan cara mencoret dan menghitamkan bulatan pada alternatif jawaban yang sesuai, serta mampu membuktikan bahwa adaptasi waktu administrasi yang telah dilakukan dalam pengerjaan tes dengan menggunakan LJK tidak mengubah hasil tes.

Hasil analisis data menunjukkan bahwa tidak terdapat perbedaan skor mentah antara kelompok kontrol dan kelompok eksperimen pada CFIT 3A ($t = 1,045$, $p > 0,05$), CFIT 3B ($t = 0,368$, $p > 0,05$), serta CFIT 3A dan 3B ($t = 0,791$, $p > 0,05$). Analisis data juga menemukan tidak terdapat perbedaan tingkat inteligensi antara kedua kelompok pada CFIT 3A ($t = 1,063$, $p > 0,05$), CFIT 3B ($t = 0,470$, $p > 0,05$), serta CFIT 3A dan 3B ($t = 0,771$, $p > 0,05$). Seluruh hipotesis penelitian yang diajukan diterima. Adaptasi waktu administrasi tidak mengubah hasil tes inteligensi.

Kelebihan yang dimiliki oleh tes yang berbasis komputer, baik menggunakan komputer dalam pengetesan maupun menggunakan lembar jawaban komputer, yaitu mampu mempercepat proses skoring dan penyajian data, bahkan data dalam jumlah yang sangat banyak. Fisteus, Pardo, dan Garcia (2013) menambahkan pembiayaan yang rendah sebagai kelebihan pengetesan berbasis komputer dengan menggunakan sistem *EyeGrade*. Sistem ini bahkan juga terbukti lebih mudah dipergunakan dan *reliabel* jika dibandingkan dengan metode yang konvensional. Hal ini akan mendukung adaptasi dari tes konvensional menuju tes berbasis komputer yang lebih baik di masa depan.

Kesimpulan

Hasil analisis data menunjukkan bahwa semua hipotesis yang diajukan dalam penelitian ini diterima. Hal ini menunjukkan bahwa adaptasi waktu standar administrasi CFIT 3 A dan 3 B berhasil dilaksanakan. Adaptasi tersebut berhasil menyamakan kesempatan untuk berpikir pada saat tes bagi subjek yang mengerjakan tes dengan cara menghitamkan bulatan dan dengan subjek yang mengerjakan tes dengan cara memberikan tanda silang pada alternatif jawaban. Namun demikian adaptasi waktu tersebut tidak mengubah atau memengaruhi hasil tes CFIT 3 A dan/atau 3 B, baik yang berupa *raw score* maupun inteligensinya.

Proses adaptasi yang dilakukan pada penelitian ini menggunakan metode eksperimen dengan dua tahap. Penggunaan desain penelitian ini berbeda dengan penelitian-penelitian sebelumnya yang meneliti adaptasi waktu dan pengetesan yang dikerjakan dengan menggunakan komputer, karena penelitian-penelitian tersebut lebih memilih untuk menggunakan *Fisher information per time unit*. Penggunaan metode eksperimen tersebut sudah tepat karena mampu memberikan hasil yang adil, baik pada pengetesan dengan menggunakan lembar jawaban manual maupun lembar jawaban komputer.

Saran

Peneliti menyarankan kepada peneliti selanjutnya atau praktisi yang ingin menggunakan lembar jawaban komputer pada saat menggunakan CFIT 3 A dan 3 B untuk menggunakan adaptasi waktu standar administrasi sesuai yang diformulasikan dalam penelitian ini. Penggunaan adaptasi waktu tersebut akan membuat hasil tes terjaga validitasnya. Hal ini sesuai

dengan tujuan dari aturan penerjemahan dan adaptasi tes yang disusun oleh ITC (2017).

Jika praktisi ingin tetap menggunakan lembar jawaban komputer LJK tanpa harus mengadaptasikan waktu administrasinya, maka disarankan untuk menggunakan LJK yang dapat dengan cara memberikan coretan pada alternatif pilihan yang tersedia. Hal ini tidak akan mengubah waktu maupun prosedur standar administrasi tes yang sesungguhnya, sehingga tidak akan memengaruhi hasil tes.

Praktisi juga dapat menggunakan tes yang langsung dikerjakan pada komputer. Alat ukur diubah bentuknya dari model cetak ke dalam model *software*. *Teste* kemudian tidak memberikan jawaban pada LJK, namun tinggal memilih atau meng-klik alternatif jawaban yang tersedia di layar komputer. Hal ini serupa dengan apa yang pernah dilakukan oleh Kumar dan Helgeson (1995). Namun isu hak cipta alat tes dan etika untuk menjaga kerahasiaan alat tes harus dipertimbangkan dengan sungguh-sungguh, karena program komputer dapat disebarluaskan secara mudah.

Kepustakaan

- Anastasi, A. & Urbina, U. (2007). *Tes psikologi* (terjemahan). Jakarta: PT. Indeks.
- Azwar, S. (2009). *Reliabilitas dan validitas*. Yogyakarta: Pustaka Pelajar.
- Azwar, S. (2010). *Pengantar psikologi inteligensi*. Yogyakarta: Pustaka Pelajar.
- Boyle, S. (1984). The effect of variations in answer-sheet format on aptitude test performance. *Journal of Occupational Psychology*, 57, 323-326.

- Brothen, T. (2012). Time limits on tests: Updating the 1-minute rule. *Teaching of Psychology*, 39(4), 288-292.
- Cattell, R. B. (1949). *Culture fair intelligence test, scale 1*. Champaign, Illinois: Institute of Personality and Ability.
- 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).
- Choe, E. M., Kern, J. L. & Chang, H. (2017). Optimizing the use of response times for item selection in computerized adaptive testing. *Journal of Educational and Behavioral Statistics*, 20(10), 1-24.
- Fisteus, J. A., Pardo, A. & Garcia, N. F. (2013). Grading multiple choice exams with low-cost and portable computer-vision techniques. *Journal of Science, Education, and Technology*, 22, 560-571.
- Goldstein, G. & Hersen, M. (2000). *Handbook of psychological assessment*. Oxford: Pergamon Press.
- Gregorie, J. & Hambleton, R. K. (2009). Advances in test adaptation research: A special issue. *International Journal of Testing*, 9, 75-77.
- Hambleton, R. K., Merenda, P. F. & Spielberger, C. D. (2005). *Adapting educational and psychological tests for cross-cultural assessment*. Hillsdale, NJ: Erlbaum.
- International Test Commission (ITC). (2017). *The ITC guidelines for translating and adapting tests (second edition)*. [www.InTestCom.org]
- Kumar, D. D. & Helgeson, S. L. (1995). Trends in computer applications in science assessment. *Journal of Science, Education, and Technology*, 4(1), 29-36.
- Miller, I. (1965). A note on the evaluation of a new answer form. *Journal of Applied Psychology*, 49, 199-201.
- Myers, A. & Hansen, C. H. (2002). *Experimental psychology*. California: Wadsworth Group.
- Saptoto, R. (2012). Perbedaan waktu pemberian jawaban dan hasil tes inteligensi ditinjau dari perbedaan lembar jawaban. *Jurnal Psikologi*, 39(2), 222-232. doi: 10.22146/jpsi.6988
- Sie, H., Finkelman, M. D., Riley, B. & Smits, N. (2015). Utilizing response times in computerized classification testing. *Applied Psychology Measurement*, 39(5), 389-405.
- SMR Indonesia. (2012, 24 Februari). *Deskripsi singkat SMR*. Diunduh 24 Februari 2012, dari SMR Indonesia Website: smrindonesia.com/index.php?option=com_content&task=view&id=45&Itemid=146.