

PENILAIAN KEPUTUSAN INVESTIGASI VARIAN: EFEK *OUTCOMES* DAN *FRAMING*

Jesica Handoko¹

Universitas Katolik Widya Mandala Surabaya
smallbink@yahoo.com

Abstract

Performance appraisal involves the judgmental evaluation, that should be made by performance appraiser on the basis of valid and reliable evidence, including when he/she evaluates his/her subordinate (a decision maker) who made variance investigation decision. It predict: (1) Outcome bias will have an impact on the perceived benefit of the investigation, (2) Investigation expenditures matched with perceived benefit are framed as costs while those without perceived benefit are framed as losses, and (3) evaluators with a cost frame provide higher performance ratings than those with a loss frame. An laboratory experiment design, with one hundred and ninety one Cost Accounting students demonstrates that outcome effect affect performance appraisal significantly. When managers did not investigate cost variance, they were evaluated more unfavorably when investigation revealed problems in the system. Further, the investigation outcome affect the perceived benefit of the investigation significantly, expenditures with perceived benefit are framed as costs while those without perceived benefit are framed as losses. Thus, paying premium frame in this research conclude moderately significant framing effect on higher performance.

Keywords: *Variance Investigation, Outcomes, Framing.*

Pendahuluan

Penilaian kinerja (*performance appraisal*) melibatkan evaluasi *judgmental* dari karakter, perilaku, atau pencapaian dari pekerja yang selanjutnya menjadi dasar untuk membuat keputusan dan rencana pengembangan personal (Kreitner dan Kinicki, 2000). Data penilaian kinerja untuk selanjutnya dapat dipakai untuk: (1) administrasi gaji, (2) umpan balik kinerja, (3) identifikasi kelemahan dan kekuatan individu, (4) dokumentasi keputusan individu, (5) pengakuan terhadap kinerja individu, (6) identifikasi kinerja yang buruk, (7) asistensi dalam identifikasi tujuan, (8) keputusan promosi, (9) retensi atau pemecatan karyawan, dan (10) evaluasi pencapaian tujuan. Oleh karena luasnya manfaat data penilaian

¹ Ucapan terima kasih ditujukan kepada Marlys G. Lipe (University of Colorado, 1993) atas ijin penggunaan instrumen risetnya, serta *experimenters*. Artikel ini pernah dipresentasikan pada Simposium Nasional Akuntansi (SNA) XI di Pontianak.

kinerja maka *judgment* penilai kinerja perlu dibuat berdasarkan bukti yang valid dan reliabel.

Disadari atau tidak, terdapat berbagai bias yang mempengaruhi *judgment* para penilai kinerja. Penelitian sekarang membahas efek bias yang ditimbulkan oleh *outcome* (*outcome bias*). *Outcome* atau *ex post information* adalah hasil akhir dari serangkaian aktivitas yang mungkin tidak terkait dengan *ex ante information* karena *outcome* tersebut tidak dapat dikendalikan oleh *decision maker* (DM). *Outcome bias* timbul saat penilaian *evaluator* dipengaruhi oleh pengetahuan tentang *outcome* (Hawkins dan Hastie, 1990; dalam Ghosh, 2005). Jika *outcome* positif maka penilai kinerja (manajemen atas) cenderung mengevaluasi bawahannya (DM) secara lebih positif dibandingkan kelayakan penilaian sesungguhnya yang didasarkan pada *ex ante information*. Sebaliknya, jika *outcome* negatif maka penilai kinerja cenderung mengevaluasi bawahannya secara lebih negatif.

Penelitian sekarang diadopsi dari Lipe (1993) yang bertujuan membuktikan pengaruh *outcome* dalam penilaian kinerja. *Setting* studi eksperimental ini tentang penilaian kinerja bawahan (*decision maker*) yang telah mengambil keputusan investigasi varian biaya produksi. Sebagai prinsip umum, suatu investigasi varian seharusnya dilakukan hanya jika manfaat yang diperkirakan lebih banyak daripada biayanya. Jika investigasi varian dilakukan, maka akan dikeluarkan sejumlah biaya dan hasil penyelidikan (*outcomes*) akan menunjukkan apakah sistem *out-of-control* (bermasalah) atau *in-control* (tidak bermasalah). Laporan biaya dan manfaat dari investigasi varian ini kemudian akan dipakai oleh manajemen atas untuk menganalisis kinerja DM. *Outcome effect* muncul jika hasil investigasi berhasil menemukan suatu masalah dalam sistem, kemudian keputusan DM akan dianggap tepat dan akan mengakibatkan pada lebih tingginya *performance rating* yang diberikan padanya dibandingkan jika hasil investigasi tidak berhasil menemukan suatu masalah dalam sistem.

Penelitian sekarang tidak hanya bertujuan menjawab rumusan masalah: (1) apakah terdapat pengaruh *outcome* terhadap penilaian kinerja, tetapi juga mencoba (2) menjelaskan bagaimana dan mengapa *outcome knowledge* mempengaruhi penilaian kinerja. Penelitian eksperimen Ghosh (2005) menemukan bahwa *perceived controllability* terhadap ukuran-ukuran *outcome* mempengaruhi *outcome effect*. Makin tingginya *perceived controllability* dari ukuran-ukuran *outcome*, maka efek *outcome* makin tinggi pula, dan sebaliknya. Penelitian Lipe (1993) menemukan suatu manfaat yang diperkirakan dari investigasi varian boleh jadi dipengaruhi oleh *outcome* investigasi. Selanjutnya, jika ditemukan sistem *out-of-control*, diduga manajemen atas (penilai kinerja) cenderung menganggap investigasi bermanfaat. Sebaliknya, jika ditemukan sistem *in-control*, maka penilai kinerja cenderung menganggap investigasi tidak bermanfaat. *Outcome effect* ini timbul karena pembingkai (*framing*) informasi yang berdampak signifikan dalam penilaian kinerja. Apabila investigasi dianggap bermanfaat maka biaya investigasi akan dipertimbangkan sebagai *cost*, sebaliknya jika tidak bermanfaat maka biaya investigasi akan dipertimbangkan sebagai *loss*. Eksperimen sekarang dilakukan untuk menjelaskan *outcome effect* dari konteks *framing*.

Bagian selanjutnya akan menjelaskan tentang tinjauan literatur dan pengembangan hipotesis. Bagian ketiga dan keempat akan mendeskripsikan metoda eksperimen dan hasil penelitian. Bagian akhir akan mendiskusikan implikasi dan keterbatasan penelitian.

Rerangka Teori dan Hipotesis

Keputusan Investigasi Varian

Biaya standar adalah biaya yang telah ditentukan sebelumnya untuk memproduksi satu unit atau sejumlah tertentu produk selama suatu periode tertentu (Carter dan Usry, 2002). Biaya standar dapat digunakan untuk: (1) menetapkan anggaran, (2) mengendalikan biaya dengan cara memotivasi karyawan dan mengukur efisiensi operasi, (3) menyederhanakan prosedur perhitungan biaya dan mempercepat laporan biaya, (4) membebaskan biaya ke persediaan bahan baku, barang dalam proses, dan barang jadi, serta (5) menetapkan tawaran kontrak dan harga jual.

Secara umum, untuk memproduksi suatu produk dibutuhkan penetapan standar fisik dan standar harga dari tiga kelompok biaya: bahan baku, tenaga kerja langsung, dan overhead pabrik. Keberhasilan penetapan standar biaya produksi ini bergantung pada keandalan, ketepatan dan penerimaan terhadap standar tersebut (Carter dan Usry, 2002). Setelah ditetapkan dan diimplementasikan, biaya standar akan dibandingkan dengan biaya aktual. Perbedaannya dianalisis dan diidentifikasi sebagai varian biaya standar. Jika biaya aktual melebihi standar, maka variannya adalah tidak menguntungkan karena kelebihan tersebut berdampak tidak menguntungkan terhadap laba. Sebaliknya, jika biaya standar melebihi aktual, maka variannya adalah menguntungkan terhadap laba.

Penelitian yang diadopsi dari Lipe (1993) ini berfokus pada keputusan investigasi varian efisiensi tenaga kerja (*labor efficiency variance/LEV*). LEV dihitung pada akhir periode pelaporan (hari, minggu, atau bulan) dengan cara membandingkan jam aktual yang digunakan dengan jam standar yang diperbolehkan untuk produksi suatu produk, yang keduanya diukur dengan tarif tenaga kerja standar (Carter dan Usry, 2002; Hansen dan Mowen, 2005). Varian yang tidak menguntungkan (jam aktual lebih besar dari jam standar) menunjukkan ketidakefisienan tenaga kerja dalam mengerjakan suatu proyek atau proses produksi. Ketidakefisienan ini dapat terjadi karena berbagai faktor, yaitu: kurangnya bahan baku, bahan baku yang bermasalah, pekerja yang tak berpengalaman, kerusakan atau keusangan mesin, perubahan metoda produksi, perencanaan dan penjadwalan yang buruk, spesifikasi disain yang buruk, ketidakpuasan kerja, dan interupsi kerja (Carter dan Usry, 2002; Hansen dan Mowen, 2005). Varian ini merupakan suatu sinyal, suatu pertanyaan yang sebaiknya diinvestigasi untuk mengetahui penyebabnya.

Keputusan investigasi varian dan mengambil tindakan perbaikan memiliki biaya yang berhubungan dengannya (Hansen dan Mowen, 2005). Sebagai prinsip umum, suatu penyelidikan seharusnya dilakukan hanya jika manfaat yang diperkirakan lebih banyak daripada biayanya. Biaya yang diperkirakan manajer (*decision maker/DM*) dalam analisis varian meliputi:

1. *Investigation expenditure* (IE) yang termasuk juga biaya dari waktu manajerial dan prosedur pengujian yang dikeluarkan.
2. Biaya memperbaiki masalah dalam sistem (membuat sistem yang *out-of-control* kembali menjadi *in-control*).
3. Biaya inkremental dari kegagalan memperbaiki sistem yang *out-of-control* atau bermasalah (Lipe, 1993).

Dalam penelitian sekarang, manajer (DM) diasumsikan telah mengestimasi IE, mempertimbangkan kemungkinan sistem *out-of-control* (berdasarkan data lampau, diskusi dengan teknisi, dan sebagainya) dan kemudian sampai pada keputusan investigasi varian. Jika investigasi varian dilakukan, maka IE akan dikeluarkan dan hasil penyelidikan (*outcomes*) akan menunjukkan apakah sistem *out-of-control* (bermasalah) atau *in-control* (tidak bermasalah). Laporan biaya dan manfaat dari investigasi varian ini kemudian akan dipakai oleh manajemen atas untuk menganalisis kinerja DM yang bertanggung jawab atas varian biaya (dalam penelitian ini, LEV).

Dampak *Outcomes* dalam Pengambilan Keputusan

Outcome adalah nilai yang dilekatkan pada suatu aktivitas. Dalam penilaian kinerja, informasi tentang *outcome* mempunyai dua efek: (a) efek pada penilaian probabilitas timbulnya *outcome* yang kemudian mempengaruhi evaluasi, dan (b) efek langsung pada penilaian kualitas keputusan (Baron dan Hersey, 1988 dalam Ghosh, 2005). Meskipun umum (normatif) bagi manajemen untuk mengevaluasi keputusan bawahannya (DM) dengan menggunakan informasi *outcome*, akan tetapi analisis semua keputusan harus dapat memisahkan antara keputusan dan *outcome*.

Sehubungan dengan keputusan investigasi varian, keputusan DM yang dibuat *ex ante* (sebelum hasil investigasi diketahui) maka seharusnya DM dievaluasi berdasarkan informasi *ex ante* (Edwards, 1984 dan Fischhoff, 1983 dalam Lipe 1993) yaitu informasi yang saat itu dibagikan di antara DM dan penilai (manajemen atas). DM yang telah memilih alternatif dengan *expected cost* terendah untuk suatu manfaat tertentu semestinya telah dapat dinilai lebih baik. *Outcome* atau *ex post information* adalah hasil akhir dari serangkaian aktivitas yang mungkin tidak terkait dengan *ex ante information* karena *outcome* tersebut tidak dapat dikendalikan oleh DM. Jadi, hanya saat makin meningkatnya *perceived controllability* dari suatu ukuran *outcome* maka tanggung jawab yang tinggi dapat dilekatkan pada konsekuensi keputusan DM (Kelley dan Michela, 1980 dalam Ghosh, 2005).

Dalam hal keputusan investigasi varian, diduga penilai kinerja (manajer atas), selain memakai *expected cost*, juga akan mempertimbangkan *outcome* investigasi (*ex post information*). Jika hasil investigasi berhasil menemukan suatu masalah dalam sistem maka keputusan DM akan dianggap tepat dan akan mengakibatkan pada lebih tingginya *performance rating* yang diberikan padanya dibandingkan jika hasil investigasi tidak berhasil menemukan suatu masalah dalam sistem. Dari penjelasan di atas akan diuji hipotesis alternatif berikut:

H1: *Outcome* dari suatu investigasi varian akan berpengaruh terhadap penilaian kinerja DM.

Perhatian terhadap *outcome bias* dalam penilaian kinerja perlu diberikan, karena akan menimbulkan berbagai masalah motivasi kerja. Jika penilai kinerja menggunakan *ex post information* untuk merevisi keputusan awalnya (*outcome bias*) dan ternyata *outcome* tersebut banyak dipengaruhi oleh banyak hal diluar kendali DM maka DM akan mengalami ketidakadilan dan kemudian muncul demotivasi kerja. Oleh karena itu, manajemen atas (penilai kinerja DM) perlu mempertimbangkan pengaruh *outcome* saat mencoba menemukan *the right balance* antara akurasi informasi dan dampaknya pada motivasi (Simons, 2000).

Framing dalam Pengambilan Keputusan

Saat terdapat kondisi ketidakpastian, pembedaan (*framing*) informasi dapat berdampak signifikan dalam pengambilan keputusan (Bazerman, 1994). Penelitian Bazerman (1983) dan Miller dan Fagley (1988) dalam Bazerman (1994) menemukan bahwa saat suatu masalah yang sama tetapi dibingkai dalam kondisi *perceived gain* dan *perceived loss* akan direspon berbeda oleh objek penelitian (pengambil keputusan). Menurut Kahneman dan Tversky (1979), dalam teori *prospectnya*, cara seseorang membingkai "masalah" secara dramatis dapat mengubah *perceived neutral point* dari suatu pertanyaan. Oleh karenanya pengambil keputusan cenderung menghindari risiko saat suatu *outcome* diekspresikan sebagai *gains* dan menghindari risiko saat suatu *outcome* diekspresikan sebagai *losses*.

Makin bervariasinya situasi ketidakpastian, *framing* suatu masalah dapat muncul dalam banyak bentuk. Kahneman dan Tversky (1979) dalam Lipe (1993) menyebutkan bahwa penggunaan *psychological* atau *mental account* untuk mengevaluasi atau mengambil keputusan juga merupakan bentuk *framing*. Penggunaan akun diskon misalnya, Mowen dan Mowen (1986) dalam Lipe (1993) menemukan bahwa pengambil keputusan cenderung mengejar diskon saat jumlah pesanan mereka kecil (dari sejumlah informasi *size of bill*).

Dalam hal investigasi varian, terdapat dua *frame* yang timbul atas biaya investigasi (IE) yaitu: IE sebagai *a cost* atau *a loss*. IE sebagai *cost* berarti IE merupakan pengeluaran, pengorbanan untuk memperoleh manfaat (Carter dan Usry, 2002). Sebaliknya, IE sebagai *loss* berarti IE merupakan pengeluaran yang tidak memperoleh manfaat. Suatu manfaat yang diperkirakan dari investigasi varian boleh jadi dipengaruhi oleh *outcome* investigasi. Pengaruh *outcome* terjadi saat ditemukannya sistem *out-of-control* (yang kemudian akan diperbaiki). Berdasarkan temuan tersebut, manajemen atas (penilai kinerja) cenderung menganggap investigasi bermanfaat. Sebaliknya, jika ditemukan sistem *in-control*, maka penilai kinerja cenderung menganggap investigasi tidak bermanfaat. Apabila investigasi dianggap bermanfaat maka IE akan dipertimbangkan sebagai *cost*, sebaliknya jika tidak bermanfaat maka IE akan dipertimbangkan sebagai *loss*. Dari penjelasan di atas akan diuji hipotesis alternatif berikut:

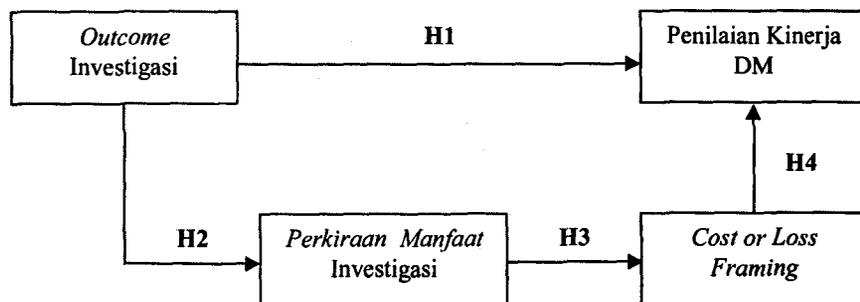
H2: *Outcome* suatu investigasi varian akan berpengaruh terhadap perkiraan manfaat investigasinya, yaitu perkiraan manfaat yang lebih besar saat sistem ditemukan *out-of-control* daripada saat sistem ditemukan *in-control*.

H3: Perkiraan manfaat dari suatu investigasi varian akan berpengaruh terhadap *framing* IE, yaitu IE akan dipertimbangkan sebagai *cost* jika IE diperkirakan menghasilkan manfaat dan IE akan dipertimbangkan sebagai *loss* jika IE diperkirakan tidak menghasilkan manfaat.

Lipe (1993) menemukan bahwa *frame* pengambil keputusan akan mempengaruhi *judgment*nya. Oleh karenanya, apabila terdapat *frame* yang berbeda atas IE investigasi varian maka penilaian kinerja yang diberikan pada DM juga akan berbeda saat IE di-*framing* sebagai *cost* (investigasi bermanfaat) daripada saat IE di-*framing* sebagai *loss*. Dari penjelasan di atas akan diuji hipotesis alternatif berikut:

H4: *Decision frame* dari penilai kinerja akan berpengaruh terhadap penilaian kinerja DM, yaitu DM akan dinilai lebih tinggi saat IE di-*framing* sebagai *cost* daripada saat IE di-*framing* sebagai *loss*.

Suatu penelitian eksperimen yang diadopsi dari Lipe (1993) akan dilakukan untuk menguji: (1) apakah *outcome* dari investigasi varian akan mempengaruhi *perceived benefit*, (2) apakah *perceived benefit* menunjukkan adanya *loss and cost frames*, serta (3) apakah *frame* mempengaruhi *judgment* penilai kinerja. Dari keempat hipotesis alternatif di atas, dapat digambarkan model penelitian yang akan diuji dalam penelitian sekarang (Gambar 1).



Gambar 1. Model Penelitian

Metoda Penelitian

Pemilihan Sampel

Sampel adalah mahasiswa Jurusan Akuntansi Fakultas Bisnis Unika Widya Mandala Surabaya yang telah mengambil mata kuliah Akuntansi Biaya. Dengan telah mengambil mata kuliah akuntansi biaya (terutama dengan topik *standard costing*) diharapkan mahasiswa telah memahami apa yang dimaksud dengan varian biaya, cara menghitung varian serta menyimpulkan apakah varian tersebut menguntungkan atau tidak menguntungkan bagi perusahaan. Dalam penelitian ini jenis varian yang digunakan adalah varian efisiensi tenaga kerja (*labor rate variance*).

Penggunaan sampel mahasiswa telah dilakukan oleh beberapa penelitian terdahulu di bidang akuntansi antara lain Cheng *et al.* (2003), Dearman dan

Shields (2005) serta Dilla dan Steinbart (2005) dalam Handoko (2006). Penggunaan sampel mahasiswa, yang berperan sebagai manajer, dipandang tepat untuk dua alasan utama (Cheng *et al.*, 2003). Pertama, semua subjek yang telah lulus mata kuliah tertentu telah familiar dengan pendekatan tradisional untuk suatu pengambilan keputusan rasional, dalam kasus penelitian sekarang yaitu penilaian keputusan investigasi varian tenaga kerja. Kedua, berbagai literatur di bidang akuntansi dan manajemen menjustifikasi kesesuaian mahasiswa sebagai *surrogates* dalam penelitian tentang pengambilan keputusan dan *judgment*. Meski merupakan *convenience sample*, akan tetapi penugasan dilakukan secara random (*randomly assignment*). Jumlah subjek untuk tiap kelompok adalah minimal 15 orang (Christensen, 1988). Partisipasi subjek bersifat sukarela.

Desain Eksperimen

Eksperimen yang akan dilakukan adalah *2x2 between-subjects design*, dengan variabel keputusan normatif (NORM) dan *outcome* investigasi varian (OUTCOME). Data yang dipresentasikan pada subjek mengindikasikan bahwa keputusan normatif adalah: (1) menginvestigasi atau (2) tidak menginvestigasi varian (menggunakan model ekspektasian biaya investigasi yang sederhana). Selanjutnya, subjek diinformasikan bahwa mereka sebagai penilai kinerja dari hasil investigasi varian DM yang menemukan bahwa sistem: (1) bermasalah (*out-of-control*) atau (2) tidak bermasalah (*in-control*). Secara keseluruhan ada 4 (empat) sel dalam eksperimen ini (Tabel 1).

Tabel 1. *2x2 Between-Subjects Experimental Design*

Outcome Investigasi	Keputusan Normatif	
	Menginvestigasi	Tidak Menginvestigasi
Sistem <i>In-Control</i> (Berjalan Baik)	Sel 1	Sel 3
Sistem <i>Out-of-Control</i> (Tidak Berjalan Baik)	Sel 2	Sel 4

Prosedur dan Skenario Eksperimen

Pengambilan data dari tiap subjek akan dilakukan dalam kelas yang diawasi oleh seorang *eksperimenter*. Cooper dan Schindler (2003) dalam Handoko (2006) menyebutkan bahwa beberapa *eksperimenter* berperan untuk mengendalikan lingkungan eksperimen dari ancaman pada validitas internal eksperimen. Apabila studi eksperimen hanya dilakukan pada oleh peneliti saja maka dikuatirkan akan muncul *demand effect*, yaitu keinginan peneliti agar subjek bertindak dan merespon sesuai yang diinginkan peneliti baik melalui cara peneliti menyapa subjek yang masuk dalam laboratorium, pemberian instruksi maupun *setting* laboratorium (Christensen, 1988).

Subjek akan berperan sebagai Direktur Produksi dari sebuah perusahaan manufaktur yang cukup besar dan menerapkan *standard costing system* dengan standar yang cukup dapat dicapai (*attainable*). Subjek membawahi beberapa manajer produksi dan tiap manajer produksi memimpin sebuah bagian

produksi/pabrik, salah satunya bernama Adi Jayadi. Berikut informasi terkait dengan kasus investigasi varian yang dilakukan oleh Adi, karena adanya varian efisiensi tenaga kerja yang cukup besar:

Adi harus memutuskan untuk menginvestigasi varian tersebut atau tidak. *Cost* investigasi diestimasikan Rp. 40.000.000. Estimasi lain menegaskan bahwa *cost* memperbaiki masalah akan sebesar Rp. 120.000.000, dan nilai sekarang (*present value*) dari *future cost* sistem yang tak terkendali dengan baik (*out-of-control*) akan sebesar Rp. 220.000.000. Data lampau menunjukkan kemungkinan sistem akan *out-of-control* (terkait dengan varian TKL yang besar) adalah 50%.

Tugas subjek penelitian adalah menilai kinerja Adi berdasarkan informasi (dengan *outcome* hasil investigasi bahwa sistem tidak berjalan baik (*out-of-control*) atau sistem berjalan baik (*in-control*)) berikut:

Anda, selaku Direktur Produksi, menerima laporan dari detil *cost* dan informasi lain terkait keputusan Adi untuk menginvestigasi varian. Anda kemudian butuh mengevaluasi kinerja Adi pada periode ini. Adi tidak bertanggung jawab atas varian bahan baku langsung atau *overhead*, tetapi keputusan investigasi dan varian tenaga kerja adalah tanggung jawab Adi.

Kinerja Adi pada periode ini akan Anda nilai sebesar (isi kotak tersedia):

0		50		100
-----	-----	-----	-----	-----
Sangat	Di bawah	Rata-rata	Di atas	Sangat
Buruk	rata-rata		rata-rata	Baik

Selanjutnya untuk menguji efek *framing* (H2-H4), subjek kemudian diminta untuk mendeskripsikan *frame* apa yang mereka pakai untuk mengevaluasi IE di antara 4 (empat) pilihan *frame*, apakah IE bermanfaat bagi perusahaan (*Benefit* (B) atau *No Benefit*(N)) dan apakah IE bersifat *Cost* (C) atau *Loss* (L).

Pengukuran dan Pengolahan Data

Variabel dependen yang diukur dari subjek adalah: (1) penilaian kinerja yang diberikan subjek pada manajer produksinya, Adi Jayadi; (2) *frame* yang dipilih subjek; (3) penilaian subjek tentang kemanfaatan IE bagi perusahaan; dan atau (4) penilaian subjek tentang IE sebagai *cost* atau *loss*.

Data (1) dipakai untuk pengujian hipotesis pertama (penilaian kinerja yang diberikan subjek pada manajer produksinya) akan dianalisis dengan *two-way analysis of variance* (ANOVA) karena alat ini dapat dipakai untuk membandingkan lebih dari dua *means* kelompok subjek serta melibatkan 2 variabel aktif (Huck, 2000). Data (2) dipakai untuk pengujian hipotesis kedua sampai keempat diuji dengan *Chi-square test* (H2-H3) ataupun dengan *one-way ANOVA* (H4). Data (3) dipakai untuk analisis tambahan untuk menguji hipotesis kedua dengan *one-way ANOVA*, sedangkan data keempat dipakai untuk pengujian H3 dan H4, dengan analisis korelasi, *Chi-square test* dan atau *t-test*.

Informasi demografis akan digunakan untuk mendeskripsikan sampel dalam eksperimen ini sedangkan hasil *manipulation checks* digunakan untuk mengetahui keseriusan atau kesiapan subjek dalam mengikuti eksperimen.

Hasil Penelitian dan Pembahasan

Manipulation Checks

Eksperimen ini menggunakan 2 jenis *manipulation checks*. Cek manipulasi pertama bertujuan mengetahui apakah subjek dapat menghitung *labor efficiency variance* (LEV) yang dialami perusahaan, sedangkan cek manipulasi kedua untuk mengetahui apakah subjek akan mendukung keputusan investigasi varian yang secara normatif (berdasarkan pertimbangan biaya-manfaat) seharusnya diinvestigasi. Berikut kedua statemen *manipulation checks*.

Selama setahun, 120.000 unit diproduksi. Jam tenaga kerja langsung aktual adalah 66.125 jam dengan biaya aktual total Rp. 648.025.000. Menurut *standard costing system* perusahaan, satu unit produk jadi membutuhkan $\frac{1}{2}$ jam TKL dengan *standard cost* Rp. 10.000 per jam.

Berapa besar varian efisiensi TKL tahun tersebut? (Berikan perhitungannya)

Adi harus memutuskan untuk menginvestigasi varian tersebut atau tidak. *Cost* investigasi diestimasikan Rp. 40.000.000. Estimasi lain menegaskan bahwa *cost* memperbaiki masalah akan sebesar Rp. 120.000.000, dan nilai sekarang (*present value*) dari *future cost* sistem yang tak terkendali dengan baik (*out-of-control*) akan sebesar Rp. 220.000.000. Data lampau menunjukkan kemungkinan sistem akan *out-of-control* (terkait dengan varian TKL yang besar) adalah 50%. Apa keputusan yang benar untuk situasi ini?

Data subjek yang salah dalam menghitung LEV atau salah membuat *judgment* investigasi tidak akan dipakai dalam analisis (Lipe, 1993). Jumlah sampel yang dapat diolah adalah 191 orang atau 70,74% dari sampel awal.

Karakteristik Demografi Sampel

Dua ratus tujuh puluh mahasiswa S-1 Fakultas Bisnis Jurusan Akuntansi berpartisipasi dalam eksperimen ini. Seluruh subjek sedang mengambil matakuliah Akuntansi Biaya, dan telah mengetahui varian biaya. Data diperoleh dari pertemuan kuliah minggu terakhir dengan seijin dosen pengajar. Dosen pengajar bertindak sebagai *eksperimenter* yang berperan untuk meminimalkan *demand effect* yang telah dijelaskan pada bagian sebelumnya. Sebanyak 79 orang (29,26%) gagal dalam menjawab pertanyaan cek manipulasi pertama atau kedua.

Tabel 2 mendeskripsikan data demografi tiap sel berdasarkan jenis kelamin sedangkan Tabel 3 menunjukkan *mean* dan *standard deviation* (*s.d*) dari umur subjek. Secara keseluruhan, terdapat 62 mahasiswa dan 129 mahasiswi berpartisipasi dalam eksperimen ini. *Mean* (deviasi standar) untuk umur subjek adalah 19,35 (0,98) tahun.

Tabel 2. Data Demografi Per Sel Berdasarkan Jenis Kelamin (n=191)

				Normative		
				Diinvestigasi	Tidak Diinvestigasi	Total
				Count	Count	Count
Outcome Berjln Baik	JnsKel Pria	12	12	24		
	Wanita	37	37	74		
Tdk berjln Baik	JnsKel Pria	19	19	38		
	Wanita	23	32	55		
Total	JnsKel Pria	31	31	62		
	Wanita	60	69	129		

Sumber: hasil pengolahan data

Tabel 3. Data Demografi Per Sel Berdasarkan Umur Subjek (n=191)

			Normative					
			Diinvestigasi		Tidak Diinvestigasi		Total	
			Mean	Std. Deviation	Mean	Std. Deviation	Mean	Std. Deviation
Outcome Berjln Baik	Umur	19.02	.81	19.48	.76	19.23	.82	
	Tdk berjln Baik	Umur	19.88	1.39	19.18	.74	19.49	1.12
Total	Umur	19.39	1.17	19.32	.76	19.35	.98	

Sumber: hasil pengolahan data

Data demografi berupa jenis kelamin dan umur subjek dianalisis lebih lanjut untuk mengetahui normalitas distribusi subjek untuk tiap sel. Hasil *chi-square test of independence* untuk jenis kelamin (JK) menunjukkan tingkat signifikansi di atas 0,05 untuk seluruh sel, sedangkan hasil tes *Kolmogorov-Smirnov* dan *Shapiro-Wilk* untuk umur (AGE) menunjukkan seluruh sel tingkat signifikansinya di bawah 0,05. Dengan kata lain data hanya menunjukkan normalitas sampel dalam penyebaran jenis kelamin. Tidak normalnya data umur disebabkan oleh beberapa subjek yang berusia di atas 20 tahun (maksimal 24, 58 tahun). Variabilitas umur subjek ini tidak terlalu luas karena hanya satu data yang menyimpang di luar 4 kali deviasi standarnya sedangkan yang lainnya masih dalam batas *interquartile range* (Huck, 2000).

Statistik Deskriptif

Tabel 4 menunjukkan statistik deskriptif variabel dependen pertama, yaitu penilaian kinerja yang diberikan subjek pada manajer produksi (Adi Jayadi) atas keputusan investigasi yang dilakukan atau tidak dilakukannya. Skala rasio yang digunakan adalah 0-100 (sangat buruk sampai sangat baik). *Mean* (deviasi standar) penilaian subjek saat manajer produksi tidak melakukan investigasi varian adalah 53,00 (25,51) ketika ditemukan bahwa sistem *in-control* dan 39,69 (21,21) ketika ditemukan bahwa sistem *out-of-control*. *Mean* (deviasi standar) penilaian subjek saat manajer produksi melakukan investigasi varian adalah 56,98 (20,71) ketika ditemukan bahwa sistem *in-control* dan 55,45 (21,36) ketika sistem *out-of-control*.

Tabel 4. Statistik Deskriptif Variabel Dependen Pertama (n=191)

			Normative					
			Diinvestigasi			Tidak Diinvestigasi		
			Mean	Std. Deviation	Count	Mean	Std. Deviation	Count
Outcome	Berjln Baik	Decision1	56.98	20.71	49	53.00	25.51	49
	Tdk berjln Baik	Decision1	55.45	21.36	42	39.69	21.21	51
	Total	Decision1	56.27	20.91	91	46.21	24.24	100

Sumber: hasil pengolahan data

Selanjutnya, separuh dari subjek (91 orang) yang termasuk dalam sel “Menginvestigasi Varian Biaya” akan dianalisis untuk mengetahui efek *framing* (Lipe, 1993). Data lengkap yang dapat diolah adalah 87 data. Dari Tabel 5 diketahui terdapat 49 data pada sel *outcome* “sistem *in-control*” dan 38 data pada sel *outcome* “sistem *out-of-control*”. Tabel 6 dan 7 menunjukkan statistik deskriptif dari data penilaian subjek tentang kemanfaatan IE bagi perusahaan dan tentang IE sebagai *cost* atau *loss* (n=184).

Tabel 5. Statistik Deskriptif Variabel Dependen Kedua (n=87)

				Count
Outcome	Berjln Baik	Decision2	Benefit, Loss	2
			Benefit, Cost	30
			No benefit, Cost	11
			No benefit, Loss	6
			Total	49
Tdk berjln Baik	Decision2	Benefit, Loss	4	
		Benefit, Cost	21	
		No benefit, Cost	6	
		No benefit, Loss	7	
		Total	38	

Sumber: hasil pengolahan data

Tabel 6. Statistik Deskriptif Variabel Dependen Ketiga (n=184)

Dependent Variable: Decision3

Normative	Outcome	Mean	Std. Deviation	N
Tidak Diinvestigasi	Tdk berjln Baik	64.68	31.232	47
	Berjln Baik	69.61	31.663	49
	Total	67.20	31.385	96
Diinvestigasi	Tdk berjln Baik	76.28	29.216	39
	Berjln Baik	68.88	27.277	49
	Total	72.16	28.231	88
Total	Tdk berjln Baik	69.94	30.712	86
	Berjln Baik	69.24	29.401	98
	Total	69.57	29.940	184

Sumber: hasil pengolahan data

Tabel 7. Statistik Deskriptif Variabel Dependen Keempat (n=184)

Dependent Variable: Decision4

Normative	Outcome	Mean	Std. Deviation	N
Tidak Diinvestigasi	Tdk berjln Baik	68.30	34.786	47
	Berjln Baik	76.94	35.500	49
	Total	72.71	35.235	96
Diinvestigasi	Tdk berjln Baik	66.87	38.499	39
	Berjln Baik	71.76	27.666	49
	Total	69.59	32.797	88
Total	Tdk berjln Baik	67.65	36.304	86
	Berjln Baik	74.35	31.767	98
	Total	71.22	34.034	184

Sumber: hasil pengolahan data

Hasil Pengujian Hipotesis

Hipotesis pertama menduga *outcome* investigasi varian akan berpengaruh terhadap penilaian kinerja DM. Hasil analisis *two-way* ANOVA dapat dilihat di Tabel 8. Variabel OUTCOME yang menunjukkan efek utama yang signifikan pada penilaian kinerja yang dilakukan subjek mendukung hipotesis pertama ($p\text{-value} = 0,023$). Penilaian kinerja ini menunjukkan bahwa manajer produksi yang tidak melakukan investigasi varian akan dinilai lebih rendah saat ditemukan sistem yang *out-of-control* (mean 39,69) dibandingkan saat ditemukan sistem *in-control* (mean 53,00).

Tabel 8. Hasil Eksperimen Efek Outcome pada Penilaian Kinerja (n=191)

Levene's Test of Equality of Error Variances

Dependent Variable: Decision1

F	df1	df2	Sig.
.576	3	187	.631

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept+NORM+OUTCOME+NORM * OUTCOME

Tests of Between-Subjects Effects

Dependent Variable: Decision1

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	9308.630 ^a	3	3102.877	6.236	.000
Intercept	499485.344	1	499485.344	1003.819	.000
NORM	4628.701	1	4628.701	9.302	.003
OUTCOME	2614.790	1	2614.790	5.255	.023
NORM * OUTCOME	1649.244	1	1649.244	3.314	.070
Error	93048.365	187	497.585		
Total	599250.000	191			
Corrected Total	102356.995	190			

a. R Squared = .091 (Adjusted R Squared = .076)

Hipotesis kedua sampai keempat (H2-H4), yang menguji efek *framing* pada penilaian kinerja, akan diuji dengan mengolah data (2). Hipotesis kedua menduga *outcome* suatu investigasi varian akan berpengaruh terhadap perkiraan manfaat investigasinya, yaitu perkiraan manfaat yang lebih besar saat sistem ditemukan *out-of-control*. Tabel 9 menunjukkan bahwa 25 dari 38 subjek (65,79%) pada sel “sistem *out-of-control*” mengindikasikan bahwa IE bermanfaat bagi perusahaan (nilai *chi-square test* 3,789 dan *p-value* = 0,05).

Tabel 9. Hasil Eksperimen Efek *Framing* pada Penilaian Kinerja (H2, n=38)

		Outcome	
		Tdk berjln Baik	Berjln Baik
		Count	Count
Dec_BNB	No Benefit	13	17
	Benefit	25	32
	Total	38	49

Test Statistics

	Dec BNB
Chi-Square ^a	3.789
df	1
Asymp. Sig.	.052

a. 0 cells (.0%) have expected frequencies less than 5. The minimum expected cell frequency is 19.0.

Hipotesis ketiga menduga perkiraan manfaat dari suatu investigasi varian akan berpengaruh terhadap *framing* IE, yaitu jika IE diperkirakan menghasilkan manfaat akan dianggap sebagai *cost* dan jika sebaliknya IE sebagai *loss*. Tabel 5 menunjukkan bahwa dari 87 subjek yang melakukan investigasi varian, sebanyak 51 orang berpendapat bahwa IE yang bermanfaat dianggap sebagai *cost* dan 13 orang berpendapat bahwa IE yang tidak bermanfaat dianggap sebagai *loss*. Tabel 10 menunjukkan bahwa 64 dari 87 subjek (73,56%) mengindikasikan bahwa IE yang bermanfaat dianggap sebagai *cost* atau IE yang tidak bermanfaat dianggap sebagai *loss* (nilai *chi-square test* 19,322 dan *p-value* = 0,000). Saat sistem ditemukan *out-of-control*, 21 dari 38 subjek menganggap investigasi varian bermanfaat dan IE dianggap sebagai *cost*.

Tabel 10. Hasil Eksperimen Efek *Framing* pada Penilaian Kinerja (H3, n=87)

Outcome	Dec_BCNL	BL or NC	Count
Berjln Baik	Dec_BCNL	BL or NC	13
		BC or NL	36
Tdk berjln Baik	Dec_BCNL	BL or NC	10
		BC or NL	28
Total	Dec_BCNL	BL or NC	23
		BC or NL	64

Test Statistics

	Dec BCNL
Chi-Square ^a	19.322
df	1
Asymp. Sig.	.000

a. 0 cells (.0%) have expected frequencies less than 5. The minimum expected cell frequency is 43.5.

Pengujian H4 tentang pengaruh *decision frame* dari penilai kinerja terhadap penilaian kinerja DM, yaitu DM akan dinilai lebih tinggi saat IE di-*framing* sebagai *cost* daripada saat IE di-*framing* sebagai *loss*. *F-test* mengindikasikan penilaian kinerja yang berbeda dari tiap kelompok *frame* subjek ($F = 1,956$; $p\text{-value}$ satu arah $< 0,10$). Penilaian kinerja untuk tiap *frame* adalah 44,61 (NL), 54,18 (NC), 52,50 (BL), dan 59,33 (BC). Lebih jauh, *mean* evaluasi subjek pada *cost and loss frame*, berturut-turut, adalah 58,61 dan 50,00 ($F = 3,593$ $p\text{-value} < 0,10$).

Analisis Tambahan

Pengukuran data variabel dependen ketiga dan keempat (Tabel 6 dan 7) digunakan untuk menguji hipotesis kedua sampai keempat. Hasil pengujian *one-way ANOVA* pada Tabel 11 menunjukkan bahwa hipotesis *null* kedua gagal ditolak ($p\text{-value} > 0,10$) meskipun subjek yang mengetahui adanya sistem yang *out-of-control*, nilainya lebih tinggi ($mean = 75,66$) dibandingkan pada kelompok sistem *in-control* ($mean = 68,88$).

Tabel 11. Hasil Eksperimen Efek *Framing* pada Penilaian Kinerja (H2, n=87)

Descriptives

SMEAN(DEC3)

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
Tdk berjin Baik	38	75.658	29.3435	4.7601	66.013	85.303	.0	100.0
Berjin Baik	49	68.878	27.2768	3.8967	61.043	76.712	.0	100.0
Total	87	71.839	28.2340	3.0270	65.822	77.857	.0	100.0

ANOVA

SMEAN(DEC3)

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	983.929	1	983.929	1.238	.269
Within Groups	67571.818	85	794.963		
Total	68555.747	86			

Data ketiga dan keempat subjek kemudian diklasifikasikan dengan kategori: nilai di bawah 50 sebagai "*no benefit*" atau "*loss*" dan nilai 50-100 sebagai "*benefit*" atau "*cost*". Hasil korelasi antara kedua variabel *perceived benefit* dan *loss/cost rating* signifikan ($r = 0,55$; $n = 191$; $p < 0,01$). Hasil uji *chi-square test* 57,855 (signifikan pada $p\text{-value} = 0,000$) mendukung hipotesis alternatif H3. Pada akhirnya, korelasi antara variabel *loss/cost rating* dan

penilaian kinerja (DEC1) yang signifikan mendukung H_4 ($r = 0,196$; $n = 191$; $p < 0,05$). Mean evaluasi subjek dengan *cost frame* dan *loss frame*, berturut-turut 53,38 dan 42,32 berbeda secara signifikan ($t = -2,751$, $p < 0,05$). Tabel 12 menunjukkan nilai korelasi antar variabel.

Tabel 12. Korelasi antara Variabel *Perceived Benefit, Loss/Cost Rating* dan Penilaian Kinerja

Correlations

		Kategori_ Dec3	Kategori_ Dec4	Decision1
Kategori_Dec3	Pearson Correlation	1	.550**	.156*
	Sig. (2-tailed)	.	.000	.031
	N	191	191	191
Kategori_Dec4	Pearson Correlation	.550**	1	.196**
	Sig. (2-tailed)	.000	.	.007
	N	191	191	191
Decision1	Pearson Correlation	.156*	.196**	1
	Sig. (2-tailed)	.031	.007	.
	N	191	191	191

** . Correlation is significant at the 0.01 level (2-tailed).

* . Correlation is significant at the 0.05 level (2-tailed).

Tabel 13. Hasil Eksperimen H4 (n=87)

Descriptives

Decision1	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
					No benefit, Loss	13		
No benefit, Cost	17	54.1176	21.37721	5.18474	43.1265	65.1088	10.00	100.00
Benefit, Cost	51	59.3333	21.40903	2.99786	53.3119	65.3547	10.00	100.00
Benefit, Loss	6	52.5000	14.74788	6.02080	37.0230	67.9770	30.00	75.00
Total	87	55.6437	20.44958	2.19243	51.2853	60.0021	10.00	100.00

ANOVA

Decision1	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	2374.279	3	791.426	1.956	.127
Within Groups	33589.675	83	404.695		
Total	35963.954	86			

Simpulan, Implikasi, dan Keterbatasan

Penelitian eksperimen dengan *2x2 mixed-subjects design* ini bertujuan untuk: (1) membuktikan pengaruh *outcome* dalam penilaian kinerja, dan (2) menjelaskan mengapa dan bagaimana *outcome knowledge* mempengaruhi evaluasi kinerja. Penelitian yang diadopsi dari studi eksperimen Lipe (1993), dengan *setting* keputusan investigasi varian pada 191 mahasiswa akuntansi, menyimpulkan bahwa:

1. Ditemukan dukungan yang signifikan secara statistis tentang efek *outcome* pada penilaian kinerja. Penilaian kinerja ini menunjukkan bahwa DM yang tidak melakukan investigasi varian akan dinilai lebih rendah saat ditemukan sistem yang *out-of-control* dibandingkan saat ditemukan sistem *in-control*. Temuan ini konsisten dengan penelitian sebelumnya (Lipe, 1993). Akan tetapi temuan lain (pada tabel 4) tidak sama dengan penelitian sebelumnya. Saat DM melakukan investigasi varian dan ditemukan sistem yang *out-of-control*, ternyata kinerjanya tidak dinilai lebih tinggi dibandingkan saat ditemukan sistem *in-control*. Penjelasan yang dapat diberikan atas temuan ini adalah subjek menganggap keputusan investigasi varian DM yang kemudian memastikan adanya sistem *in-control* merupakan tindakan yang baik. Bazerman (1994) menyimpulkannya sebagai "*framing of paying premium*", yaitu IE tertentu yang dikeluarkan dalam investigasi varian dianggap sebagai "premi asuransi" untuk mengamankan manfaat tertentu.
2. Ditemukan dukungan yang signifikan secara statistis tentang efek *framing* pada penilaian kinerja. Apabila ditemukan sistem yang *out-of-control*, IE dianggap bermanfaat bagi perusahaan, sehingga konsisten dengan Lipe (1993). Hasil penelitian sekarang juga konsisten dengan Bazerman (1994) tentang "*framing of paying premium*", yaitu saat ditemukan sistem yang *in-*

control, IE dianggap bermanfaat bagi perusahaan. Selanjutnya, ditemukan jika IE menghasilkan manfaat akan dianggap sebagai *cost* dan jika sebaliknya IE sebagai *loss*. Akhirnya, ditemukan pengaruh *decision frame* terhadap penilaian kinerja DM, yaitu DM akan dinilai lebih tinggi saat IE di-*framing* sebagai *cost* daripada saat IE di-*framing* sebagai *loss*, akan tetapi temuan ini hanya *moderately significant*.

Temuan penelitian ini berimplikasi pada pentingnya mempertimbangkan efek *outcome* pada penilaian kinerja, serta apa yang menyebabkan timbulnya *outcome effect* tersebut yang diantaranya adalah cara pandang pengambil keputusan. Keterbatasan penelitian ini diduga mempengaruhi generalisasi hasil penelitian. Pertama, penelitian ini mengutamakan validitas internal dengan menggunakan responden mahasiswa sebagai proksi pengambil keputusan. Alasan digunakannya mahasiswa akuntansi adalah kemampuan teoritis mereka untuk menganalisis data yang diberikan, yang terkait dengan bahasan *standard costing*. Penggunaan sampel mahasiswa juga telah banyak dilakukan. Kedua, penelitian ini dilakukan dengan persetujuan dosen pengajar matakuliah bukan inisiatif pribadi dari mahasiswa. Dampak yang muncul adalah banyaknya jumlah sampel yang salah mengisi *manipulation checks* (yaitu sebanyak 79 orang atau 29,26% sampel awal). Diduga mereka kurang memberi perhatian pada kasus yang diberikan, sehingga respon mereka tidak dapat diolah. Meskipun tidak mendapatkan *random sample*, tetapi penugasan subjek ke tiap *treatment cells* telah dilakukan secara random, sesuai saran Cowles (1974) dalam Christensen (1988).

Berdasarkan keterbatasan yang ada, penelitian selanjutnya diharapkan memperbaiki atau memperluas *setting* studi eksperimen. Penggunaan subjek manajer sesungguhnya, variasi varian biaya serta variabel-variabel lain yang mempengaruhi penilaian kinerja DM (mempertimbangkan makin banyaknya industri yang menerapkan *standard cost*) akan memperkuat generalisasi temuan. Hal ini penting dilakukan mengingat pertimbangan keputusan investigasi varian dan penilaian atas hasil investigasi tersebut merupakan konsekuensi alamiah dari penerapan *standard costing*.

Daftar Rujukan

- Bazerman, M.H. 1994. *Judgment in Managerial Decision Making*. 3rd edition. John Wiley & Sons, Inc.
- Carter, W.K. dan M.F. Usry. 2002. *Cost Accounting*. 13rd edition. Dame, a division of Thomson Learning.
- Cheng, M.M., A.K.D. Schulz, P.F. Lockett, dan P. Booth. 2003. The Effects of Hurdle Rates on the Level of Escalation of Commitment in Capital Budgeting. *Behavioral Research in Accounting*. 15: pp. 63-85.
- Christensen, L.B. 1988. *Experimental Methodology*. 4th Edition. Allyn and Bacon, Inc.
- Ghosh, D. 2005. Alternative Measures of Managers' Performance, Controllability, and the Outcome Effect. *Behavioral Research in Accounting*. 17: pp. 55-70.

- Handoko, J. 2006. Pengaruh Emosi Negatif dalam Pemilihan Alternatif Investasi Modal: Perbandingan Keputusan Individu dan Kelompok. *Jurnal Riset Akuntansi*. 10(3): pp. 303-320.
- Hansen, D.R. dan M.M. Mowen. 2005. *Management Accounting*. 7th edition. South-Western of Thomson Learning.
- Huck, S.W. 2000. *Reading Statistics and Research*. 3th Edition. Addison Wesley Longman, Inc.
- Kahneman, D. dan A. Tversky. 1979. Prospect Theory: An Analysis of Decision Under Risk. *Econometrica*. 47: pp. 263-291.
- Kreitner, R. dan A. Kinicki. 2000. *Organizational Behavior*. 5th edition. McGraw Hill Companies, Inc.
- Simons, R. 2000. *Performance Measurement and Control Systems for Implementing Strategy*. Prentice Hall-New Jersey.