

Studi Eksperimental: Mengurangi Bias Pengukuran Umum Balanced Scorecard Dalam Penilaian Kinerja Pada Mahasiswa Program Manajemen Bisnis

Krisanti Pramono dan Juniarti

Akuntansi Bisnis Universitas Kristen Petra

Email: yunie@peter.petra.ac.id

ABSTRAK

Penelitian ini bertujuan untuk menggabungkan kedua pendekatan dari penelitian Roberts, Albright, & Hibbets (2004) yang menggunakan pendekatan *disaggregated/ mechanically aggregated* dan penelitian Dilla & Steinbart (2005) yang menggunakan pendekatan pengetahuan agar mengurangi bias pengukuran umum yang ada dalam evaluasi kinerja menggunakan Balanced Scorecard, dan melihat pengaruh antara evaluasi kinerja dengan alokasi kompensasi. Uji hipotesis penelitian ini menggunakan analisis *repeated measures ANOVA*, *independent sample t-test* serta analisis regresi linier berganda. Hasil penelitian menunjukkan bahwa pendekatan *disaggregated/ mechanically aggregated* memiliki pengaruh secara signifikan dalam mengurangi bias pengukuran umum pada evaluasi kinerja menggunakan Balanced Scorecard, namun pendekatan pengetahuan terbukti tidak berpengaruh dalam mengurangi bias tersebut. Selain itu, ditemukan adanya pengaruh yang signifikan antara evaluasi kinerja terhadap alokasi kompensasi. Evaluasi kinerja menggunakan Balanced Scorecard menjelaskan hampir semua dari variasi yang ada dalam keputusan kompensasi.

Kata kunci: Balanced Scorecard, pengukuran umum, pengukuran unik, bias pengukuran umum, *debiasing*, *disaggregated/mechanically aggregated*, pengetahuan, alokasi kompensasi

ABSTRACT

This study aims to combine the two approaches from Roberts, Albright, & Hibbets's (2004) study that used disaggregated/mechanically aggregated approach and Dilla & Steinbart's (2005) study that used training approach to reduce the common measure bias in the performance evaluation using Balanced Scorecard, and to investigate the influence between performance evaluation and compensation allocation. This study's hypothetical test uses repeated measures ANOVA, independent sample t-test and multiple regression linear analysis. Results indicate that disaggregated/mechanically aggregated approach had significant influence in overcoming the common measure bias, but the knowledge approach was proven not helping in reducing the bias. Furthermore, it was found that there was a significant influence between performance evaluation and compensation allocation. Balanced Scorecard performance evaluations explained almost all of the variations in compensation decisions.

Keywords: *Balanced Scorecard, common measures, unique measures, common measure bias, debiasing, disaggregated/mechanically aggregated, knowledge, bonus allocation*

PENDAHULUAN

Balanced Scorecard didefinisikan sebagai sistem pengukuran yang secara relatif kompleks dan mahal (Lipe & Salterio, 2000; Dilla & Steinbart, 2005). Salah satu elemen biayanya adalah manajemen waktu (Lipe & Salterio, 2000). Balanced Scorecard harus dimanfaatkan secara optimal agar investasi tersebut tidak sia – sia, terutama agar dapat diaplikasikan kepada kepentingan perusahaan (Kootanae, Kootanae, Hoseinian, & Talari 2013) dan searah dengan tujuan perusahaan (Malina & Selto, 2001). Dalam mengaitkan penilaian Balanced Scorecard dengan kompensasi, individual dalam perusahaan tidak akan menerima kompensasi bila kinerja dalam periode yang telah ditentukan tidak sesuai dengan yang diharapkan, sehingga memotivasi karyawan untuk mencapai kinerja yang lebih seimbang dan sesuai dengan tujuan jangka pendek maupun jangka panjang (Kaplan & Norton, 2007).

Namun pada penelitian – penelitian sebelumnya (Lipe & Salterio, 1998, 2000; Ittner, Larcker, & Meyer, 2003; Roberts, Albright, & Hibbets, 2004; Dilla & Steinbart, 2005; Banker, Chang, & Pizzini, 2009) ditemukan bahwa partisipan yang diminta mengambil peran sebagai evaluator untuk menilai kinerja, cenderung lebih mengandalkan pengukuran umum (*common measures*) daripada pengukuran unik (*unique measures*) ketika mengambil keputusan. Lipe & Salterio (1998, 2000) menemukan kecenderungan ini dalam penelitian mereka, yang disebut sebagai bias pengukuran umum (*common measure bias*) (Grevinga, 2013).

Melanjutkan penelitian Lipe & Salterio (2000), Roberts, Albright, & Hibbets (2004) meneliti sebab munculnya bias pengukuran umum dan mencari pendekatan yang berpotensi untuk mengurangi bias tersebut. Hasilnya, bias pengukuran umum dapat diminimalisir meski masih ditemukan adanya ketidakkonsistenan dalam pemberian kompensasi (Grevinga, 2013).

Sementara itu, Dilla & Steinbart (2005) menemukan bahwa pengambil keputusan yang sudah menjalani pelatihan dalam teori dan pengembangan Balanced Scorecard menggunakan baik pengukuran umum maupun unik untuk mengevaluasi

kinerja dan mengalokasikan kompensasi secara cukup optimal, namun tetap menaruh fokus lebih besar kepada pengukuran umum daripada pengukuran unik ketika membuat pertimbangan.

Berdasarkan penelitian sebelumnya (Roberts, Albright, & Hibbets 2004; Dilla & Steinbart, 2005), motivasi peneliti untuk melanjutkan penelitian tersebut adalah untuk mengurangi bias yang ada pada penelitian – penelitian sebelumnya dengan menggabungkan pendekatan yaitu *disaggregated/mechanically aggregated* untuk mengurangi bias yang ada dalam penilaian kinerja menggunakan Balanced Scorecard. Berdasarkan pernyataan di atas, peneliti merumuskan masalah sebagai berikut:

1. Apakah dengan menggunakan pendekatan *disaggregated/mechanically aggregated* dan pengetahuan dapat mengurangi bias yang ada dalam evaluasi kinerja menggunakan Balanced Scorecard?
2. Apakah penilaian kinerja berbasis Balanced Scorecard memunculkan pengaruh terhadap alokasi kompensasi?

BALANCED SCORECARD

Dalam artikel *Harvard Business Review*, Kaplan & Norton (1992) mendefinisikan Balanced Scorecard sebagai satu set pengukuran yang memberikan pandangan bisnis secara keseluruhan dengan cepat dan komprehensif kepada manajer tingkat atas. Balanced Scorecard dikembangkan karena manajer tidak hanya ingin dinilai berdasarkan pengukuran tradisional yaitu keuangan saja, namun juga non-keuangan yaitu pengukuran kinerja (Grevinga, 2013). Pengukuran tersebut adalah *learning and growth, internal business process, customer perspective* (pengukuran non-keuangan) dan *financial performance* (pengukuran keuangan).

Pengukuran Umum dan Unik

Pengukuran umum (*common measures*) adalah pengukuran baik keuangan maupun non-keuangan yang berlaku di seluruh perusahaan. Grevinga (2013) mendefinisikan pengukuran umum sebagai

pengukuran yang sesuai dengan strategi organisasional, sementara pengukuran unik (*unique measures*) adalah pengukuran yang disesuaikan dengan strategi divisional. Pengukuran unik adalah pengukuran baik keuangan maupun non-keuangan yang berlaku hanya pada masing – masing divisi.

Bias Pengukuran Umum (*Common Measure Bias*)

Lipe & Salterio (1998) adalah orang pertama yang menekankan keberadaan dari bias pengukuran umum (*common measure bias*) dalam Balanced Scorecard (Grevinga, 2013). Bias pengukuran umum adalah keengganan dari pembuat keputusan untuk memasukkan informasi unik (*unique information*) karena informasi ini membutuhkan usaha kognitif yang lebih untuk diproses (Lipe & Salterio, 2000; Grevinga, 2013), yang menyebabkan evaluator cenderung menggunakan strategi penyederhanaan terhadap penilaian kinerja (Lipe & Salterio, 1998).

Tidak dilibatkannya pengukuran unik dalam pengambilan keputusan dapat membuat strategi per divisi tidak terefleksikan dalam keputusan tersebut, sehingga menghilangkan manfaat potensial dari Balanced Scorecard sebagai sistem manajemen strategi (Libby, Salterio, & Webb, 2002; Dilla & Steinbart, 2005; Roberts, Albright, & Hibbets, 2004; Lipe & Salterio, 2000).

Mengurangi Bias Pengukuran Umum

Dengan adanya bias pengukuran umum, peneliti berusaha mengurangi bias tersebut dengan melakukan *debiasing*. Penelitian ini menggunakan pendekatan *disaggregated/ mechanically aggregated* dan pengetahuan untuk mengurangi bias pengukuran umum yang ada dalam penilaian kinerja menggunakan Balanced Scorecard.

Hasil penelitian Roberts, Albright, & Hibbets (2004) menunjukkan bukti bahwa pendekatan *disaggregated/mechanically aggregated* efektif dalam mengurangi bias pengukuran umum yang ditemukan oleh Lipe & Salterio (2000) ketika Balanced Scorecard

digunakan untuk mengevaluasi kinerja secara holistik. Menyediakan *disaggregated* Balanced Scorecard dapat meningkatkan fokus dan usaha kognitif secara keseluruhan ketika mereka mengevaluasi menggunakan semua pengukuran. Berdasarkan hasil penelitian tersebut, maka diajukan hipotesis sebagai berikut.

H₁: Pendekatan *disaggregated / mechanically aggregated* dapat membantu mengurangi bias pengukuran umum yang ada dalam evaluasi kinerja menggunakan Balanced Scorecard.

Dilla & Steinbart (2005) menginvestigasi pengaruh dari hal pengetahuan dalam evaluasi kinerja. Hasilnya mereka menemukan bahwa dengan menyediakan informasi tambahan mengenai Balanced Scorecard, dapat meningkatkan konsistensi dalam menentukan pertimbangan dengan memudahkan dalam membandingkan pengukuran individual Balanced Scorecard antar divisi (Dilla & Steinbart, 2005; Grevinga, 2013). Berdasarkan hasil penelitian-penelitian tersebut, maka diajukan hipotesis sebagai berikut.

H₂: Pengetahuan dapat membantu mengurangi bias pengukuran umum yang ada dalam evaluasi kinerja menggunakan Balanced Scorecard.

Pengukuran Balanced Scorecard dan Alokasi Kompensasi

Salah satu jenis keputusan kompensasi melibatkan alokasi bonus (Dilla & Steinbart, 2005). Kinerja dan alokasi kompensasi sangat berkorelasi (Robert, Albright, & Hibbets, 2004, Lipe & Salterio, 2000) karena kompensasi dapat memotivasi karyawan untuk bekerja lebih baik dan meningkatkan kinerja mereka. Kaplan & Norton (1996, 2001) menyarankan kepada organisasi maupun perusahaan untuk meningkatkan manfaat dari Balanced Scorecard dengan cara mengaitkan hasil penilaian kinerja tersebut terhadap keputusan kompensasi. Keputusan kompensasi dipengaruhi oleh hasil evaluasi kinerja dari manajer divisi (Grevinga, 2013; Roberts, Albright, & Hibbets 2004). Berda-

sarkan hasil penelitian-penelitian tersebut, maka diajukan hipotesis sebagai berikut.

H₃: Evaluasi kinerja menggunakan Balanced Scorecard berpengaruh terhadap alokasi kompensasi.

Kajian Penelitian Terdahulu

Lipe & Salterio (2000) menginvestigasi apakah pengukuran umum mendominasi evaluasi kinerja menggunakan Balanced Scorecard. Pengambil keputusan cenderung lebih menggunakan pengukuran umum daripada pengukuran unik ketika dihadapkan pada keduanya. Lipe & Salterio (2000) menduga bahwa evaluasi kinerja menggunakan Balanced Scorecard akan dipengaruhi baik oleh pengukuran umum maupun unik.

Roberts, Albright, & Hibbets (2004) melanjutkan penelitian Lipe & Salterio (2000). Tujuan dari penelitian ini adalah untuk meneliti pendekatan yang berpotensi dalam mengurangi bias yang ada dalam penilaian kinerja Balanced Scorecard. Penelitian Roberts, Albright, & Hibbets (2004) membuktikan bahwa pendekatan *disaggregated/mechanically aggregated* efektif dalam mengurangi bias pengukuran umum ketika Balanced Scorecard digunakan untuk menilai kinerja, meski penelitian tersebut masih menimbulkan ketidak-konsistenan dalam mengaplikasikan Balanced Scorecard sebagai standar alokasi kompensasi (Grevinga, 2013).

Dilla & Steinbart (2005) juga melanjutkan penelitian Lipe & Salterio (2000) namun menggunakan pendekatan yang berbeda, yaitu pengetahuan. Informasi atau pengetahuan tambahan dapat meningkatkan konsistensi dalam pertimbangan keputusan (Dilla & Steinbart, 2005; Grevinga, 2013). Pengambil keputusan dengan pengetahuan yang cukup akan menyertakan baik pengukuran umum maupun unik ketika disuruh untuk mengevaluasi kinerja, dan dapat mengalokasikan kompensasi secara cukup optimal ketika alokasi kompensasi dihubungkan dengan penilaian kinerja (Grevinga, 2013), namun tetap lebih mengandalkan pengukuran umum daripada pengukuran unik.

METODOLOGI PENELITIAN

Penelitian ini menggunakan pendekatan eksperimen laboratorium dengan alasan peneliti ingin mendapatkan hasil yang sesuai dengan tujuan penelitian, sehingga peneliti melakukan manipulasi terhadap satu atau lebih variabel independen atau faktor yang merupakan fenomena yang menyebabkan munculnya hasil yang diteliti (Nazir, 1988). Partisipan dalam penelitian ini adalah 40 orang mahasiswa yang sedang menjalani S1 fakultas ekonomi jurusan Manajemen Bisnis Universitas Kristen Petra dengan syarat sedang/sudah mengikuti mata kuliah Akuntansi Manajemen yang memiliki sedikit materi teoritis mengenai Balanced Scorecard.

Desain Eksperimen

Uji eksperimen penelitian mereplika eksperimen yang dilakukan oleh Roberts, Albright, & Hibbets (2004), namun dengan menambahkan pendekatan pengetahuan yang dilakukan oleh Dilla & Steinbart (2005). Desain eksperimen dalam penelitian ini diadopsi dari penelitian Lipe & Salterio (2000), Roberts, Albright, & Hibbets (2004), serta Dilla & Steinbart (2005), yaitu menggunakan desain 2 x 2 x 2, terdiri dari 2 faktor *between-subjects* (Umum dan Unik) dan 1 faktor *within-subjects* (Divisi). Faktor *between-subjects* adalah faktor yang membandingkan antara subyek dengan subyek, sementara faktor *within-subjects* adalah faktor yang membandingkan kinerja per divisi secara keseluruhan.

Sebelum uji eksperimen dimulai, 50% dari jumlah keseluruhan partisipan diminta untuk mengikuti *review* singkat mengenai Balanced Scorecard, sementara 50% sisanya tidak dipersilakan masuk ataupun mendengarkan *review* tersebut. Setelah *review* dilakukan, Partisipan diminta berperan sebagai eksekutif senior sebuah perusahaan yaitu WCS Incorporated yang bertugas menilai kinerja 2 manajer setiap divisi. 2 divisi yaitu RadWear (RAD) yang khusus menjual pakaian remaja dan WorkWear (WORK) yang khusus menjual seragam kerja disajikan dalam kasus beserta visi dan strategi tiap divisi; misi dari perusahaan WCS; dan instruksi tentang bagaimana mengerjakan kasus tersebut.

Partisipan diberi waktu untuk membaca kasus selama 15-20 menit.

Setelah partisipan selesai membaca kasus, peneliti memberi presentasi mengenai sejarah dan latar belakang perusahaan untuk memperluas informasi yang dimiliki oleh partisipan mengenai perusahaan WCS, kemudian membagikan Balanced Scorecard tiap divisi dan meminta partisipan untuk menyelesaikan 2 langkah *disaggregated*, dimana mereka 1) mengukur tiap kinerja manajer dengan setiap 16 *item* dari Balanced Scorecard menggunakan skala dari 0 (*Unacceptable*) sampai dengan 100 (*Excellent*), kemudian 2) mengalikan pertimbangan individual tersebut dengan bobot yang telah ditentukan, dan menjumlahkan hasil skornya untuk mendapatkan skor *aggregate* (total skor) pada tiap divisi. Bobot yang telah ditentukan untuk Unique adalah 64% dari total bobot.

Variabel dependen pertama dalam penelitian ini adalah evaluasi kinerja terhadap 2 manajer divisi. Variabel dependen kedua dalam penelitian ini adalah alokasi kompensasi sebesar seratus juta rupiah (Rp 100.000.000,00). Partisipan diminta untuk mengalokasikan kompensasi tersebut kepada 2 manajer divisi tanpa ketentuan dasar untuk menentukan alokasi tersebut.

HASIL PENELITIAN DAN PEMBAHASAN

Profil Partisipan

Tabel 1. Jenis Kelamin Berdasarkan Partisipan Dengan Pengetahuan dan Tanpa Pengetahuan

			Frequency	Percent	Valid Percent	Cumulative Percent
Pengetahuan	Valid	Pria	8	40,0	40,0	40,0
		Wanita	12	60,0	60,0	100,0
	Total	20	100,0	100,0		
Tanpa Pengetahuan	Valid	Pria	11	55,0	55,0	55,0
		Wanita	9	45,0	45,0	100,0
	Total	20	100,0	100,0		

Tabel 2. Statistif Deskriptif Umum dan Unik

Pengetahuan	N	Minimum	Maximum	Mean	Std. Deviation
Dengan Pengetahuan Umum	20	52.03	61.68	56.902	2.312
Dengan Pengetahuan Unik	20	70.23	91.45	87.238	4.631
Tanpa Pengetahuan Umum	20	49.52	60.60	56.044	3.107
Tanpa Pengetahuan Unik	20	79.36	93.46	85.935	4.441

Uji Hipotesis 1

Tabel 3. Analisis Metode *Repeated Measures* ANOVA dari Partisipan Keseluruhan

Source	Type III Sum of Squares	Df	Mean Square	F	Sig.
Intercept	102329.745	1	102329.745	40589.631	.000
Umum	136.725	1	136.725	54.233	.000
Unik	4534.190	1	4534.190	1798.510	.000
Umum * Unik	632.559	1	632.559	250.908	.000
Error	393.289	156	2.521		

Source	Divisi	Type III Sum of Squares	Df	Mean Square	F	Sig.
Divisi	Linear	4534.190	1	4534.190	1798.510	.000
Divisi * Umum	Linear	632.559	1	632.559	250.908	.000
Divisi * Unik	Linear	102329.745	1	102329.745	40589.631	.000
Divisi * Umum * Unik	Linear	136.725	1	136.725	54.233	.000
Error(Divisi)	Linear	393.289	156	2.521		

Tabel 3 menunjukkan interaksi antara Divisi dan Umum (Divisi*Umum) menghasilkan nilai F sebesar 250,908 dengan signifikansi sebesar $0,000 < 0,05$. Interaksi antara Divisi dan Unik (Divisi*Unik) menghasilkan nilai F sebesar 40589,631 dengan signifikansi sebesar $0,000 < 0,05$, sehingga dapat disimpulkan bahwa hasil Unik pada divisi RadWear dan divisi WorkWear berbeda secara signifikan. Berdasarkan hasil ini, dapat disimpulkan bahwa evaluasi kinerja menggunakan pendekatan *dissaggregated/mechanically aggregated* berhasil dalam mengurangi bias pengukuran umum yang ada dalam evaluasi kinerja menggunakan Balanced Scorecard oleh partisipan secara keseluruhan.

Tabel 4. Analisis Repeated Measures ANOVA Partisipan Dengan Pengetahuan

Tests of Between-Subjects Effects

Measure: MEASURE_1
Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	51940.489	1	51940.489	23254.499	.000
Umum	49.807	1	49.807	22.299	.000
Unik	2300.758	1	2300.758	1030.082	.000
Umum * Unik	312.174	1	312.174	139.765	.000
Error	169.751	76	2.234		

Tests of Within-Subjects Contrasts

Measure: MEASURE_1

Source	Divisi	Type III Sum of Squares	df	Mean Square	F	Sig.
Divisi	Linear	2300.758	1	2300.758	1030.082	.000
Divisi * Umum	Linear	312.174	1	312.174	139.765	.000
Divisi * Unik	Linear	51940.489	1	51940.489	23254.499	.000
Divisi * Umum * Unik	Linear	49.807	1	49.807	22.299	.000
Error(Divisi)	Linear	169.751	76	2.234		

Tabel 4 menunjukkan interaksi antara Divisi dan Umum (Divisi*Umum) menghasilkan nilai F sebesar 139,765 dengan signifikansi sebesar $0,000 < 0,05$. Interaksi antara Divisi dan Unik (Divisi*Unik) menghasilkan nilai F sebesar 23254,499 dengan signifikansi sebesar $0,000 < 0,05$, sehingga dapat disimpulkan bahwa hasil Unik pada divisi RadWear dan divisi WorkWear berbeda secara signifikan, sehingga dapat disimpulkan bahwa evaluasi kinerja menggunakan pendekatan *dissagregated/ mechanically aggregated* berhasil dalam mengurangi bias pengukuran umum yang ada dalam evaluasi kinerja menggunakan Balanced Scorecard pada partisipan yang diberi pengetahuan tentang Balanced Scorecard.

Tabel 5. Analisis Repeated Measures ANOVA Partisipan Tanpa Pengetahuan

Tests of Between-Subjects Effects

Measure: MEASURE_1
Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	50395.091	1	50395.091	17855.934	.000
Umum	89.850	1	89.850	31.836	.000
Unik	2233.680	1	2233.680	791.435	.000
Umum * Unik	320.413	1	320.413	113.528	.000
Error	214.496	76	2.822		

Tests of Within-Subjects Contrasts

Measure: MEASURE_1

Source	Divisi	Type III Sum of Squares	df	Mean Square	F	Sig.
Divisi	Linear	2233.680	1	2233.680	791.435	.000
Divisi * Umum	Linear	320.413	1	320.413	113.528	.000
Divisi * Unik	Linear	50395.091	1	50395.091	17855.934	.000
Divisi * Umum * Unik	Linear	89.850	1	89.850	31.836	.000
Error(Divisi)	Linear	214.496	76	2.822		

Tabel 5 menunjukkan interaksi antara Divisi dan Umum (Divisi*Umum) menghasilkan nilai F sebesar 113,528 dengan signifikansi sebesar $0,000 < 0,05$. Selain itu, interaksi antara Divisi dan Unik (Divisi*Unik) menghasilkan nilai F sebesar 17855,934 dengan signifikansi sebesar $0,000 < 0,05$, sehingga dapat disimpulkan bahwa hasil Unik pada divisi RadWear dan divisi WorkWear berbeda secara signifikan, sehingga dapat disimpulkan bahwa evaluasi kinerja menggunakan pendekatan *dissagregated/ mechanically aggregated* berhasil dalam mengurangi bias pengukuran umum yang ada dalam evaluasi kinerja menggunakan Balanced Scorecard pada partisipan yang tidak diberi pengetahuan tentang Balanced Scorecard.

Uji Hipotesis 2

Tabel 6. Independent Sample T-Test

Group Statistics

CODE	N	Mean	Std. Deviation	Std. Error Mean
Umum Pengetahuan	20	56.902	2.312	.517
Tanpa Pengetahuan	20	56.044	3.107	.695
Unik Pengetahuan	20	87.2380	4.63134	1.03560
Tanpa Pengetahuan	20	85.9350	4.44130	.99311

Independent Samples Test

	Levene's Test for Equality of Variances		t-test for Equality of Means							
	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference		
								Lower	Upper	
Umum	Equal variances assumed	1,046	,313	,990	38	,328	,858	,866	-.696	2.611
	Equal variances not assumed			,990	35,109	,329	,858	,866	-.900	2.615
Unik	Equal variances assumed	,874	,356	,908	38	,370	1,303	1,435	-1,602	4,208
	Equal variances not assumed			,908	37,933	,370	1,303	1,435	-1,602	4,208

Hasil *independent sample t-test* menyimpulkan bahwa tidak terdapat perbedaan signifikan hasil pengukuran antara partisipan yang diberi pengetahuan tentang Balanced Scorecard dan partisipan yang tidak diberi pengetahuan tentang

Balanced Scorecard, baik Umum maupun Unik. Berdasarkan hasil ini, dapat disimpulkan bahwa pendekatan pengetahuan yang memadai tentang Balanced Scorecard tidak mengurangi bias pengukuran umum yang ada dalam evaluasi kinerja menggunakan Balanced Scorecard.

Uji Hipotesis 3

Tabel 7. Uji F Regresi

Tabel 4.23 Uji F Regresi

ANOVA ^a						
Model	Sum of Squares	Df	Mean Square	F	Sig.	
1	Regression	.006	4	.002	3424.354	.000 ^b
	Residual	.000	30	.000		
	Total	.006	34			

a. Predictors: (Constant), WorkUnik, RadUnik, WorkUmum, RadUmum
 b. Dependent Variable: Y_Ratio

Dari Tabel 7, diketahui bahwa uji F menghasilkan nilai signifikansi sebesar 0,000 < 0,05, maka disimpulkan bahwa RadUmum, RadUnik, WorkUmum dan WorkUnik secara simultan berpengaruh signifikan terhadap alokasi kompensasi.

Tabel 8. Nilai R Square Regresi

Model Summary^b

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.999 ^a	.998	.998	.0006850

a. Predictors: (Constant), WorkUnik, RadUnik, WorkUmum, RadUmum
 b. Dependent Variable: Y_Ratio

Dari Tabel 8 diketahui Adjusted R Square sebesar 0,998 menunjukkan bahwa alokasi kompensasi dipengaruhi oleh RadUmum, RadUnik, WorkUmum dan WorkUnik sebesar 99,8%, sedangkan sisanya yaitu 0,2% dipengaruhi oleh faktor lain.

Hasil penelitian membuktikan bahwa pendekatan *disaggregated/ mechanically aggregated* dapat mengurangi bias pengukuran umum karena kedua pengukuran digunakan secara proporsional. Hasil ini konsisten dengan penelitian sebelumnya yaitu penelitian Roberts, Albright, & Hibbets (2004) yang menggunakan pendekatan yang sama.

Penelitian Dilla & Steinbart (2005) membuktikan bahwa pelatihan dapat mengurangi bias pengukuran umum dalam evaluasi kinerja menggunakan Balanced Scorecard. Dalam penelitian ini, bias pengukuran umum terbukti sudah berkurang, namun tidak ada perbedaan yang signifikan

antara kelompok yang diberi pengetahuan tentang Balanced Scorecard dan kelompok yang tidak diberi pengetahuan, memberikan kesimpulan bahwa pengetahuan tidak berpengaruh secara signifikan terhadap berkurangnya bias pengukuran umum. Dermer (1973, dalam Debusk, Killough, & Brown, 2005) mengungkapkan bahwa karakteristik kognitif dari seorang pengguna informasi dapat mempengaruhi persepsi pengguna tersebut tentang mana informasi yang penting dan mana yang tidak, sehingga dapat berdampak pada bagaimana informasi tersebut mempengaruhi perilaku terakhirnya. Ketika partisipan memiliki intoleransi terhadap ketidakpastian yang tinggi, mereka merasa bahwa informasi yang dimiliki sudah cukup untuk melakukan evaluasi kinerja, oleh karena itu tidak begitu memerdulikan informasi tambahan yang diberikan oleh peneliti karena partisipan lebih memercayai penilaian subjektifitas mereka dalam mengevaluasi kinerja (Debusk, Killough, & Brown, 2005). Selain itu, uji eksperimen diadakan sehari sebelum dan pada minggu terakhir ujian akhir universitas, sehingga kemungkinan besar partisipan tidak fokus mengerjakan evaluasi kinerja karena faktor sibuk belajar atau kelelahan menghadapi ujian.

Dalam mengaitkan penilaian Balanced Scorecard dengan kompensasi, individual dalam perusahaan tidak akan menerima kompensasi bila kinerja dalam periode yang telah ditentukan tidak sesuai dengan yang diharapkan. Ketentuan ini dapat memotivasi karyawan untuk mencapai kinerja yang lebih seimbang dan sesuai dengan tujuan jangka pendek maupun jangka panjang (Kaplan & Norton, 2007). Hasil penelitian ini konsisten dengan penelitian sebelumnya (Lipe & Salterio, 2000; Roberts, Albright, & Hibbets, 2004; Dilla & Steinbart, 2005). Dalam penelitian ini evaluasi kinerja terbukti berpengaruh ketika partisipan diminta untuk mengalokasikan kompensasi.

KESIMPULAN

Berdasarkan hasil pengujian statistik yang telah dilakukan terhadap 40 mahasiswa program Manajemen Bisnis Fakultas Ekonomi dalam eksperimen yang diadakan

oleh peneliti, maka dapat diambil kesimpulan sebagai berikut:

1. Pendekatan *disaggregated/mechanically aggregated* terbukti signifikan dalam mengurangi bias pengukuran umum yang ada dalam evaluasi kinerja menggunakan Balanced Scorecard. Hal ini dikarenakan pendekatan ini mengurangi usaha kognitif dalam mengevaluasi kinerja yang dilakukan oleh partisipan, sehingga disimpulkan H_1 diterima.
2. Pendekatan pengetahuan terbukti tidak membantu dalam mengurangi bias pengukuran umum yang ada dalam evaluasi kinerja menggunakan Balanced Scorecard. Hal ini diduga karena keterbatasan jumlah sampel yang ada, tingkat kelelahan dari partisipan yang disebabkan oleh ujian akhir yang diadakan pada minggu yang sama, dan karakteristik dari partisipan yang cenderung tidak begitu memerdulikan informasi tambahan yang diberikan, sehingga disimpulkan H_2 ditolak.
3. Evaluasi kinerja menggunakan Balanced Scorecard terbukti berpengaruh terhadap alokasi kompensasi, dengan tujuan untuk meningkatkan manfaat dari Balanced Scorecard sendiri, dan meningkatkan motivasi dari karyawan agar menghasilkan kinerja yang sesuai dengan tujuan baik jangka pendek maupun jangka panjang, sehingga disimpulkan H_3 diterima.

Keterbatasan

Partisipan dalam uji eksperimen penelitian ini merupakan mahasiswa program Manajemen Bisnis Universitas Kristen Petra yang minimal sedang mengambil mata kuliah Akuntansi Manajemen yang memberikan sedikit teori mengenai Balanced Scorecard, sehingga pada dasarnya memiliki sedikit pengetahuan teoritis mengenai Balanced Scorecard meskipun tidak memiliki pengalaman menggunakan Balanced Scorecard. Tidak ada nilai minimal yang menjadi syarat partisipan untuk mengikuti eksperimen ini

Penelitian ini tidak dapat mengukur seberapa besar pengetahuan tentang Balanced Scorecard yang dimiliki oleh partisipan dengan pengetahuan sebelum dan sesudah pendekatan pengetahuan diberikan. Eksperimen dilakukan pada saat yang kurang

tepat, yaitu pada minggu UAS sehingga partisipan sulit memfokuskan diri pada eksperimen karena eksperimen diadakan sebelum dan saat UAS.

Saran

Penelitian selanjutnya diharapkan untuk menambah syarat dalam mendapatkan partisipan yaitu lulus dengan minimal B pada mata kuliah yang menjelaskan Balanced Scorecard secara teoritis, sehingga memiliki kemungkinan yang lebih tinggi dalam memahami baik teori maupun penerapan Balanced Scorecard. Penelitian selanjutnya disarankan untuk mengukur pengetahuan yang dimiliki oleh partisipan sebelum dan sesudah pendekatan pengetahuan, dengan memberikan pertanyaan-pertanyaan terkait dengan teori dan penerapan mengenai Balanced Scorecard.

Penelitian selanjutnya diharapkan untuk memilih saat yang tepat untuk melakukan eksperimen. Sebaiknya eksperimen selanjutnya dilakukan pada saat minggu tenang atau saat hari biasa agar partisipan tidak terganggu oleh adanya ujian akhir yang dapat mengurangi fokus partisipan ketika mengevaluasi kinerja.

DAFTAR PUSTAKA

- Debusk, G.K., Killough, L.N., & Brown, R.M. (2005). Intolerance of Ambiguity Effects on Evaluations with the Balanced Scorecard. *The Innovation Congress: Issues In Innovation*, vol. 3, issue 1.
- Dilla, W. N., & Steinbart, P. J., (2005). Relative Weighting of Common and Unique Balanced Scorecard Measures by Knowledgeable Decision Makers. *Behavioral Research in Accounting*, 17, 43-53.
- Grevinga, K. H. M. (2013). Common Measure Bias in the Balanced Scorecard: an Experiment with Undergraduate Students. Unpublished undergraduate thesis, University of Twente, Enschede.

- Ittner, C. D., Larcker, D. F. & Meyer, M. W., "Subjectivity and the Weighting of Performance Measures: Evidence from a Balanced Scorecard." SSRN Working Paper, 2003.
- Kaplan, R. S., & Norton, D. P. (1992). *The Balanced Scorecard*. Retrieved from *Harvard Business Review website*: <http://www.hbr.org/>
- Kaplan, R. S., & Norton, D. P. (1996, Fall). *The Balanced Scorecard: Translating Strategy into Action*. Boston: Harvard Business School Press.
- Kaplan, R. S., & Norton, D. P. (2001). *The Strategy-Focused Organization: How Balanced Scorecard Companies Thrive in the New Business Environment*. Boston: Harvard Business School Press.
- Kaplan, R. S., & Norton, D. P. (2007, July-August). Using the Balanced Scorecard as a Strategic Management System. *Harvard Business Review: Best of HBR*.
- Kootanae, A. J., Kootanae, H. J., Hoseinian, H., & Talari, H. F. (2013). The Balanced Scorecard, Alphabet of the Modern Management: From Concept to Implement. *Advanced in Management & Applied Economics*, 3 (1), 47-60.
- Libby, T., Salterio, S., & Webb, A. (2002, June). The Balanced Scorecard: The Effects of Assurance and Process Accountability on Managerial Judgment. Retrieved from <http://ssrn.com/abstract=530803>
- Lipe, M. G., & Salterio S. E. (1998). *The Balanced Scorecard: Judgemental Effects of Information Organization and Diversity*. SSRN Working Paper, November 1998.
- Lipe, M. G., & Salterio S. E. (2000, July). The Balanced Scorecard: Judgemental Effects of Common and Unique Performance Measure. *The Accounting Review*, 75 (3), 283- 298.
- Malina, M. A., & Selto, F. H. (2001). Communicating and Controlling Strategy: an Empirical Study of the Effectiveness of the Balanced Scorecard. *Journal of Management Accounting Research*, 13, 47-90.
- Nazir, M. (1988). *Metode Penelitian* (edisi 3). Jakarta: Ghalia Indonesia.
- Roberts, M. L., Albright, T. L., & Hibbets, A. R. (2004). Debiasing Balanced Scorecard Evaluations. *Behavioral Research in Accounting*, 16 (1), 75-88.