

# STUDI EKSPERIMENTAL: MENGURANGI BIAS PENGUKURAN UMUM BALANCE SCORECARD DALAM PENILAIAN KINERJA PADA MAHASISWA SI PROGRAM MANAJEMEN PARIWISATA

Janet Georgina dan Juniarti  
Akuntansi Bisnis Universitas Kristen Petra  
Email: yunie@petra.ac.id

## ABSTRAK

Hasil studi Lipe & Salterio (2000) menemukan adanya dominasi pengukuran umum pada hasil evaluasi kinerja menggunakan BSC, dan hal ini mengakibatkan bias pengukuran umum. Berbagai pendekatan dilakukan untuk mengatasi bias pengukuran umum, namun belum secara optimal mengurangi bias pengukuran umum.

Penelitian ini menggabungkan dua pendekatan yaitu *disaggregate/ mechanically aggregate* dan pengetahuan untuk mengurangi bias pengukuran umum. Penelitian ini juga menguji apakah evaluasi kinerja menggunakan BSC mempengaruhi alokasi kompensasi. Penelitian ini menggunakan metode penelitian eksperimental. Hipotesis pertama diuji dengan menggunakan *repeated measures ANOVA*, hipotesis kedua diuji menggunakan *independent sample T-Test* dan hipotesis kedua diuji menggunakan regresi linier berganda. Pendekatan *disaggregate/ mechanically aggregate* dan pengetahuan berhasil mengurangi bias pengukuran umum dan hasil evaluasi kinerja BSC mempengaruhi alokasi kompensasi.

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

## ABSTRACT

The study Lipe & Salterio (2000) found a common measurement dominance on the results of the performance evaluation using the BSC, and this has resulted in a common measure bias. Various approaches are taken to address the common measure bias, but was not optimally reduced the common measure bias.

This study combined two approaches which were *disaggregate / mechanically aggregate* and knowledge to reduce the common measure bias. The study also examined whether the performance evaluation using the BSC affected the allocation of compensation. This study used an experimental research method. The first hypothesis was tested by using *repeated measures ANOVA*, the second hypothesis was tested by using *independent sample t-test* and the second hypothesis was tested by using *multiple linear regression*. The *Disaggregate /mechanically aggregate* and knowledge approaches were managed to reduce the common measure bias of BSC and the results of BSC performance evaluation influenced the allocation of compensation.

Keywords; Balanced Scorecard, common measure, unique measure, *debiasing*, common measure bias, *disaggregated/ mechanically aggregated*, knowledge, allocation of compensation

## PENDAHULUAN

Kekuatan penting yang dimiliki BSC adalah masing-masing unit bisnis dalam

perusahaan akan memiliki *scorecard* yang secara spesial dibuat untuk unit bisnis tersebut (Kaplan & Norton, 1996b). Masing-masing unit bisnis dalam perusahaan

mengembangkan pengukuran BSC miliknya untuk merefleksikan tujuan dan strategi spesifik unit bisnisnya, dan untuk menunjukkan apa yang harus dicapai oleh unit bisnis untuk mencapai kesuksesan yang selanjutnya akan membantu perusahaan mencapai tujuannya (Lipe & Salterio, 2000; Robert, Albright & Hibbets, 2004; Dilla & Steinbart, 2005; Kaplan & Norton, 1996, 2000, 2001; Libby, Salterio & Webb, 2002).

Namun, hasil studi Lipe & Salterio (2000) menemukan bahwa, evaluasi kinerja unit yang dibuat oleh manajer menggunakan informasi yang disajikan dalam format BSC banyak dipengaruhi oleh kinerja yang diukur secara umum untuk semua unit dan sedikit dipengaruhi oleh ukuran kinerja unik dari masing-masing unit bisnis. Dominasi ukuran umum ini mengakibatkan munculnya bias pengukuran umum (Slovic & MacPhillamy, 1974).

Dominasi pengukuran umum pada hasil evaluasi kinerja menggunakan BSC akan menjadi masalah. Karena, dengan hanya berfokus pada pengukuran umum dan mengabaikan pengukuran unik akan mengurangi manfaat dari BSC sebagai sistem manajemen strategi perusahaan (Libby, et. al. 2002; Lipe & Salterio, 2000; Dilla & Steinbart, 2005). Manfaat BSC berkurang karena semua *item* dalam BSC merupakan ukuran yang penting dari kinerja strategi perusahaan (Robert, et. al. 2004) yang membantu perusahaan mencapai tujuannya. Hal ini sejalan dengan Holmstrom (1979) yaitu pengukuran kinerja akan bernilai hanya jika pengukuran tersebut menyediakan informasi yang unik.

Manajer yang hanya fokus pada pengukuran umum dan mengabaikan pengukuran unik ketika mengevaluasi kinerja bawahan atau karyawan, juga akan mempengaruhi kinerja karyawan. Dimana, karyawan akan mengurangi usaha mereka terhadap aktivitas-aktivitas yang dianggap tidak diutamakan, karena aktivitas-aktivitas tersebut dianggap tidak akan mempengaruhi hasil evaluasi kinerjanya dan kompensasi yang akan di terima (Holmstrom & Milgrom, 1991; Malina & Selto, 2001). Hal ini akan menghambat perusahaan mencapai tujuannya karena seperti yang dikemukakan oleh Kaplan dan Norton (1996) semua ukuran dalam BSC dibuat untuk mencapai tujuan perusahaan.

Selain itu, mengabaikan ukuran unik tiap unit bisnis dalam pengukuran kinerja akan mengakibatkan manajer mengambil

keputusan yang kurang tepat, salah satunya keputusan kompensasi. Hal ini karena keputusan agen dipengaruhi oleh hal-hal yang dimasukkan ke dalam evaluasi kinerja dan kompensasi (Holmstrom & Milgrom 1991). Pengambilan keputusan yang kurang tepat mengenai kompensasi berdasarkan hasil evaluasi kinerja BSC, akan menciptakan ketidakpuasan terhadap BSC (Dilla & Steinbart, 2005). Karena, karyawan yang melihat evaluasi kinerja yang tidak konsisten akan merasa tidak adil (Gibbs et al. 2004).

Karena semua ukuran yang termasuk dalam BSC dianggap relevan untuk mencapai tujuan perusahaan (Kaplan & Norton, 1996, 2001), banyak pendekatan dilakukan untuk mengurangi bias pengukuran umum pada BSC. Seperti yang dilakukan oleh Robert, et. al. (2004) yaitu dengan menggunakan pendekatan *disaggregated/mechanically aggregated* dan Dilla & Steinbart (2005) memberikan pelatihan yang cukup bagi partisipan dan mengadakan sesi praktik mendesain dan menerapkan BSC. Kedua pendekatan tersebut telah berhasil mengurangi bias pengukuran umum dalam evaluasi kinerja menggunakan BSC, namun hasilnya bias belum berkurang secara optimal. Indikasi dari kurang optimalnya hasil dari kedua pendekatan tersebut adalah, karena kedua pendekatan tersebut dilakukan secara parsial atau terpisah.

Kurang *optimal-nya* hasil dari kedua pendekatan sebelumnya baik menggunakan *disaggregated/mechanically aggregated BSC* maupun memberikan pengetahuan dan pelatihan yang cukup mengenai BSC dalam mengurangi bias pengukuran umum, memberikan daya tarik bagi peneliti untuk melakukan penelitian lebih lanjut yaitu dengan menggabungkan kedua metode tersebut yang sebelumnya dilakukan secara terpisah.

### **Pengertian Balance Scorecard**

BSC merupakan alat pengukuran kinerja yang komprehensif, karena terdiri dari pengukuran financial dan non financial (tradisional dan non tradisional), pengukuran non financial meliputi tiga perspektif yaitu *learning and growth, internal business process, dan customer perspective* (Kaplan & Norton, 1992, 1996, 2001). Kombinasi dari keempat ukuran kinerja keuangan dan non-keuangan tersebut dihubungkan dengan visi, misi dan strategi perusahaan (Grevinga, 2013). BSC juga merupakan alat pengukuran kinerja yang

digunakan untuk menerjemahkan strategi kedalam ukuran-ukuran yang lebih jelas baik ukuran financial dan ukuran non financial (Kaplan & Norton, 2001).

BSC juga merupakan sistem akuntansi manajemen yang memungkinkan perusahaan untuk menerjemahkan visi, misi dan strategi kedalam tindakan (Agrawal, 2008). BSC membantu manajemen dalam mengkomunikasikan tujuan strategis perusahaan dengan secara jelas menghubungkan tujuan tersebut dengan target-target kinerja yang relevan dan mengevaluasi kinerja karyawan berdasarkan hal tersebut (Johnson et al. 2013).

Tujuan utama BSC adalah untuk mengidentifikasi, mengkomunikasikan, dan mengimplementasikan strategi perusahaan di seluruh lapisan perusahaan (Kaplan & Norton 1992,1996).

### **Pengukuran Umum dan Pengukuran Unik**

Pengukuran umum adalah pengukuran kinerja baik financial maupun non financial yang berlaku untuk seluruh lapisan perusahaan. Pada umumnya ukuran umum merupakan ukuran yang diterapkan di seluruh unit dalam perusahaan, (Kaskey, 2006). Hal ini dikarenakan unit bisnis merupakan bagian dari perusahaan, maka beberapa ukuran akan berlaku umum untuk semua unit bisnis dan diidentifikasi sebagai ukuran umum (Gagne, 2006). Karena unit bisnis merupakan bagian dari perusahaan maka BSC yang digunakan unit harus konsisten dengan strategi perusahaan secara umum (Kaskey, 2006). Contoh pengukuran umum antara lain, dari perspektif *financial* adalah *return on sales*, dari perspektif *customer* adalah *repeat sales*, dari perspektif *internal business process* adalah *return on supplier*, dari perspektif *learning and growth* adalah *hours of employee training per employee* (Lipe & Salterio, 2000).

Kekuatan penting dari BSC adalah bahwa setiap unit bisnis dalam organisasi akan memiliki *scorecard* sendiri secara khusus dirancang untuk unit tersebut (Kaplan & Norton, 1996). Beberapa dari pengukuran spesifik dipilih untuk tiap unit bisnis dalam perusahaan akan berbeda satu dengan lainnya, karena pengukuran tersebut harus dibuat berdasarkan tujuan spesifik dan strategi masing-masing unit (Lipe & Salterio, 2000). Sehingga unit-unit pada tingkatan yang

sama dalam organisasi akan memiliki beberapa pengukuran kinerja yang umum disamping pengukuran kinerja yang unik yang sesuai dengan bisnis dan strategi unit-unit tersebut (Lipe & Salterio, 2000).

Menggunakan ukuran unik untuk mengukur kesuksesan unit bisnis merupakan tujuan diciptakannya BSC (Kaplan & Norton, 1996).

### **Bias Pengukuran Umum**

Bias pengukuran umum sering dijelaskan sebagai ketidakmampuan pengambil keputusan untuk menyertakan informasi unik dalam evaluasi kinerja, karena informasi ini membutuhkan usaha cognitive yang lebih untuk bisa diproses (Slovic & MacPhillamy 1974; Lipe & Salterio, 2005). Sejalan dengan Johnson et al. (2013) yang menyatakan bahwa ketergantungan pengambil keputusan pada pada ukuran-ukuran umum disebabkan oleh ketidakmampuan untuk mempertimbangkan informasi-informasi yang independen. Hal ini terjadi karena individu dibatasi oleh ketersediaan informasi dalam memori mereka (Tversky dan Kahneman 1973 dan 1974). Dalam melakukan evaluasi kinerja, manajer menggunakan kedua pengukuran yaitu pengukuran umum dan unik untuk menilai kinerja bawahan atau unit bisnis yang ada dalam perusahaan. Namun semua BSC cenderung menggunakan pengukuran umum tertentu (Kaplan & Norton, 1996) hal ini akhirnya menyebabkan munculnya bias pengukuran umum.

Dominasi pengukuran umum dalam evaluasi perbandingan didasari setidaknya tiga faktor. Pertama, pengukuran umum membentuk himpunan informasi yang lebih sederhana dari keseluruhan informasi, dan secara kognitif lebih mudah untuk menguasai dan mengelola sedikit dari pada banyak informasi (Anderson, 1990). Kedua, tidak hanya mengurangi informasi secara keseluruhan tetapi juga dapat mengakibatkan berkurangnya kategori atau jenis informasi yang harus di kelola (Lipe & Salterio, 2002). Ketiga, pengukuran umum merupakan satu-satunya informasi yang tersedia untuk secara langsung digunakan untuk membandingkan hasil kinerja (Robert, Albright & Hibbets, 2004).

### **Mengurangi Bias Pengukuran Umum (Debiasing)**

Menurut Bazerman (2006) *Debiasing* didefinisikan sebagai prosedur untuk mengurangi atau mengeliminasi bias dari strategi kognitif pengambil keputusan.

Menurut Servas (2010) ada beberapa level teknik dasar untuk mengatasi bias, level yang pertama adalah *knowledge, experience, and expertise*, level yang kedua adalah *incentive and accountability* dan level ketiga adalah *training dan tools*.

*Disaggregated / Mechanically Aggregated* merupakan alat untuk mengatasi bias dengan cara menyederhanakan informasi yang perlu diproses (Robert, 2004). Pendekatan *disaggregated/mechanically aggregated* dalam hal ini juga diartikan sebagai, partisipan yang berperan sebagai manajer dalam eksperimen ini tidak melakukan penilaian dengan membandingkan kinerja satu unit dengan unit lain tapi dengan melakukan penilaian secara individual untuk masing-masing unit (Grevinga, 2013).

Pendekatan lainnya yang dilakukan untuk menangkal bias adalah dengan pengetahuan (*knowledge*) seperti yang dijelaskan dalam penelitian Servas (2010). Hal ini sejalan dengan pernyataan bahwa pengetahuan yang lebih luas akan menghasilkan penilaian yang lebih baik di berbagai macam tugas audit (Bonner 1990; Bonner dan Lewis 1990). Pengetahuan yang luas akan meningkatkan kemampuan seseorang untuk fokus dalam penelitian, karena meningkatkan pemahaman orang tersebut (Nelson, 1982). Hal ini terutama relevan dengan penilaian yang didasarkan pada BSC, semakin baik pembuat keputusan memahami teori dan struktur BSC seharusnya semakin mudah untuk menggabungkan ukuran-ukuran baik umum dan unik ketika membandingkan dan mengevaluasi kinerja masing-masing unit bisnis (Dilla & Steinbart, 2005).

### **Pengukuran Kinerja menggunakan BSC dan Alokasi Kompensasi**

Secara konseptual, perusahaan harus menghubungkan kompensasi dengan BSC untuk meningkatkan manfaat dari BSC (Kaplan & Norton 1996, 2001). Sehingga manfaat dari BSC akan bertambah, yang sebelumnya hanya menjadi alat untuk implementasi dan memonitor strategi serta sebagai alat untuk mengukur kinerja (Kaplan & Norton 1996), akan bertambah manfaatnya yaitu memberikan pertimbangan bagi manajer dalam hal keputusan kompensasi.

Holmstrom & Milgrom (1991) menemukan bahwa keputusan agen dalam hal ini manajemen dipengaruhi oleh hal-hal yang dimasukkan ke dalam evaluasi kinerja, sehingga mengabaikan ukuran unik tiap unit bisnis dalam pengukuran kinerja akan mengakibatkan manajer mengambil keputusan yang kurang tepat, contohnya keputusan mengenai kompensasi. Pengambilan keputusan yang kurang tepat mengenai kompensasi berdasarkan hasil evaluasi kinerja BSC, akan menciptakan ketidakpuasan terhadap BSC (Dilla & Steinbart, 2005). Karena, karyawan yang melihat evaluasi kinerja yang tidak konsisten akan merasa tidak adil (Gibbs at al. 2004). Ketidakpuasan ini dapat menyebabkan perusahaan berhenti menghubungkan BSC dengan keputusan kompensasi (Ittner, Larcker & Meyer, 2003).

### **METODE PENELITIAN**

Penelitian ini menggunakan metode penelitian eksperimental. Hipotesis pertama diuji dengan menggunakan *repeated measures ANOVA*, hipotesis kedua diuji menggunakan *independent sample T-Test* dan hipotesis kedua diuji menggunakan regresi linier berganda.

Variabel dependent pertama adalah perbedaan evaluasi kinerja antara dua divisi. Score diberikan partisipan kepada kedua manajer berdasarkan pengukuran partisipan atas kinerja manajer. Variabel dependen kedua adalah alokasi kompensasi. Variabel dependen ini digunakan untuk menguji apakah alokasi kompensasi dipengaruhi oleh evaluasi kinerja berdasarkan BSC atau tidak.

Hipotesis 1 dari penelitian diuji dengan menggunakan *repeated measures ANOVA*. Hipotesis yang akan diuji adalah apakah pendekatan *disaggregated/mechanically aggregated* dapat mengurangi bias pengukuran umum yang ada pada BSC. Hipotesis kedua dari penelitian diuji dengan menggunakan *independent sample T-Test*. Hipotesis yang akan diuji adalah apakah pendekatan pengetahuan yang memadai tentang BSC akan mengurangi bias ukuran umum. Hipotesis ketiga ini diuji dengan menggunakan analisis regresi. Dengan hipotesis apakah evaluasi kinerja menggunakan BSC akan mempengaruhi alokasi kompensasi bagi manajer kedua divisi.

### **HASIL PENELITIAN DAN PEMBAHASAN**

Partisipan dalam eksperimen ini adalah mahasiswa S1 Fakultas Ekonomi jurusan pariwisata sebanyak 30 mahasiswa yang minimal sedang mengikuti matakuliah akuntansi manajemen, sehingga diharapkan partisipan telah memahami teori BSC dan contoh-contoh penerapan BSC secara teoritis. Data diambil di akhir masa perkuliahan semester genap 2014.

Untuk mengetahui apakah pendekatan *disaggregated/mechanically aggregated* dapat mengurangi bias ukuran umum pada evaluasi kinerja menggunakan BSC, maka digunakan metode *repeated measures ANOVA*. Dengan melihat interaksi antara ukuran umum dan unik dengan kedua divisi maka akan terlihat apakah pendekatan ini berhasil mengurangi bias pengukuran umum. Berikut ini disajikan tabel hasil uji dengan menggunakan *repeated measures ANOVA* untuk keseluruhan data baik yang mendapatkan pengetahuan maupun tidak.

Tabel 4.2. *Repeated Measures ANOVA Keseluruhan*

**Tests of Between-Subjects Effects**

Measure: MEASURE\_1  
Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	77616.067	1	77616.067	9822.445	.000
Common	.645	1	.645	.082	.776
Unique	22.915	1	22.915	2.900	.091
Common * Unique	7409.037	1	7409.037	937.626	.000
Error	916.621	116	7.902		

**Tests of Within-Subjects Contrasts**

Measure: MEASURE\_1

Source	division	Type III Sum of Squares	df	Mean Square	F	Sig.
division	Linear	22.915	1	22.915	2.900	.091
division * Common	Linear	7409.037	1	7409.037	937.626	.000
division * Unique	Linear	77616.067	1	77616.067	9822.445	.000
division * Common * Unique	Linear	.645	1	.645	.082	.776
Error(division)	Linear	916.621	116	7.902		

Berdasarkan tabel *between-subjects effect* terlihat bahwa interaksi dari pengukuran umum dan unik menunjukkan pengaruh yang signifikan ( $F = 937.626$ ;  $p\text{-value} < 0.05$ ), hal ini menunjukkan bahwa kedua pengukuran baik umum dan unik berpengaruh terhadap evaluasi kinerja manajer. Pada tabel *within-subjects effect* dapat diketahui bahwa terdapat pengaruh

yang signifikan ketika ukuran umum diinteraksikan dengan divisi ( $F = 937.626$ ;  $p\text{-value} < 0.05$ ), begitu pula ketika ukuran unik diinteraksikan dengan divisi ( $F = 322.445$ ;  $p\text{-value} < 0.05$ ).

Berikut ini disajikan tabel hasil uji dengan menggunakan *repeated measures ANOVA* untuk partisipan yang mendapatkan pengetahuan.

Tabel 4.3. *Repeated Measures ANOVA Pengetahuan*

**Tests of Between-Subjects Effects<sup>a</sup>**

Measure: MEASURE\_1  
Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	45415.752	1	45415.752	20928.918	.000
Common	6.599	1	6.599	3.041	.087
Unique	7.105	1	7.105	3.274	.076
Common * Unique	4199.413	1	4199.413	1935.214	.000
Error	121.520	56	2.170		

a. Kode = Pengetahuan

**Tests of Within-Subjects Contrasts<sup>a</sup>**

Measure: MEASURE\_1

Source	division	Type III Sum of Squares	df	Mean Square	F	Sig.
division	Linear	7.105	1	7.105	3.274	.076
division * Common	Linear	4199.413	1	4199.413	1935.214	.000
division * Unique	Linear	45415.752	1	45415.752	20928.918	.000
division * Common * Unique	Linear	6.599	1	6.599	3.041	.087
Error(division)	Linear	121.520	56	2.170		

a. Kode = Pengetahuan

Dalam kelompok partisipan yang diberikan pengetahuan interaksi antara ukuran umum dan ukuran unik menunjukkan pengaruh yang signifikan ( $F = 1935.214$ ;  $p\text{-value} < 0.05$ ), hal ini menunjukkan bahwa penggunaan kedua pengukuran ini yaitu umum dan unik berpengaruh terhadap evaluasi kinerja manajer. Sedangkan pada tabel *within-subjects effect* dapat diketahui bahwa terdapat pengaruh yang signifikan secara statistik pada interaksi ukuran umum dan divisi ( $F = 1935.214$ ;  $p\text{-value} < 0.05$ ) demikian pula halnya dengan pengaruh pada interaksi antara ukuran unik dan divisi ( $F = 20928.918$ ;  $p\text{-value} < 0.05$ ) juga menunjukkan effect yang signifikan.

Berikut ini disajikan tabel hasil uji dengan menggunakan *repeated measures ANOVA* untuk yang tidak mendapatkan pengetahuan.

Tabel 4.4. *Repeated Measures ANOVA Pengetahuan*

Tests of Between-Subjects Effects<sup>a</sup>

Measure: MEASURE\_1

Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	32719.519	1	32719.519	8173.753	.000
Common	13.723	1	13.723	3.428	.069
Unique	16.845	1	16.845	4.208	.045
Common * Unique	3240.641	1	3240.641	809.554	.000
Error	224.168	56	4.003		

a. Kode = Non Pengetahuan

Tests of Within-Subjects Contrasts<sup>a</sup>

Measure: MEASURE\_1

Source	division	Type III Sum of Squares	df	Mean Square	F	Sig.
division	Linear	16.845	1	16.845	4.208	.045
division * Common	Linear	3240.641	1	3240.641	809.554	.000
division * Unique	Linear	32719.519	1	32719.519	8173.753	.000
division * Common * Unique	Linear	13.723	1	13.723	3.428	.069
Error(division)	Linear	224.168	56	4.003		

a. Kode = Non Pengetahuan

Pada kelompok partisipan yang tidak diberikan pengetahuan, interaksi dari pengukuran umum dan unik menunjukkan pengaruh yang signifikan ( $F = 809.554$ ;  $p\text{-value} < 0.05$ ), hal ini menunjukkan bahwa penggunaan kedua pengukuran ini yaitu umum dan unik berpengaruh terhadap evaluasi kinerja manajer. Tabel *within-subjects effect* menunjukkan terdapat pengaruh yang signifikan secara statistik pada interaksi ukuran umum dan divisi ( $F = 809.554$ ;  $p\text{-value} < 0.05$ ) demikian pula halnya dengan pengaruh pada interaksi antara ukuran unik dan divisi ( $F = 8173.753$ ;  $p\text{-value} < 0.05$ ) juga menunjukkan effect yang signifikan.

Jika *independent sample t test* menghasilkan nilai signifikansi  $< 0.05$  ( $\alpha=5\%$ ), maka dapat disimpulkan bahwa terdapat perbedaan yang signifikan antara evaluasi pengukuran dari partisipan dengan pengetahuan BSC dan tanpa pengetahuan BSC. Berikut adalah hasil *independent sample t test with equal variances assumed* data evaluasi pengukuran umum dan unik pada kelompok partisipan dengan pengetahuan BSC dan tanpa pengetahuan BSC:

Tabel 4.5. Hasil *Independent sample T-Test*

Group Statistics

Kode	N	Mean	Std. Deviation	Std. Error Mean
Common Pengetahuan	15	54.1540	3.21248	.82946
Common Non Pengetahuan	15	45.2633	3.58847	.92654
Unique Pengetahuan	15	101.4793	2.47332	.63861
Unique Non Pengetahuan	15	86.8367	5.65267	1.45951

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
Common	Equal variances assumed	.610	.441	7.149	28	.000	8.89067	1.24358	6.34332	11.43802
	Equal variances not assumed			7.149	27.664	.000	8.89067	1.24358	6.34192	11.43841
Unique	Equal variances assumed	6.930	.014	9.191	28	.000	14.64267	1.59311	11.37933	17.90601
	Equal variances not assumed			9.191	19.171	.000	14.64267	1.59311	11.31026	17.97507

Berdasarkan Tabel di atas diketahui bahwa rata-rata evaluasi kinerja manajer dengan pengukuran umum oleh partisipan dengan pengetahuan BSC adalah sebesar 54.154 sedangkan oleh partisipan tanpa pengetahuan BSC sebesar 45.263. Rata-rata nilai evaluasi kinerja dengan pengukuran unik sebesar 101.479 oleh partisipan dengan pengetahuan BSC dan 86.837 oleh partisipan tanpa pengetahuan BSC. Uji beda nilai evaluasi dengan pengukuran umum dan unik menghasilkan nilai signifikansi masing-masing sebesar 0.000 (sangat kecil) dimana nilai ini lebih kecil dari 0.05 ( $\alpha = 5\%$ ).

Analisis regresi dilakukan dengan variabel bebas yaitu Rc, Ru, Wc dan Wu serta variabel terikat yaitu alokasi bonus (kompensasi). Berikut ini adalah hasil analisis regresi:

Tabel 4.6. Nilai R Square

Model Summary <sup>b</sup>				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.785 <sup>a</sup>	.615	.554	.16616

a. Predictors: (Constant), Wu, Wc, Ru, Rc

b. Dependent Variable: Yrasio

Koefisien determinasi (*R Square*) adalah nilai yang digunakan untuk mengetahui pengaruh dari variabel bebas terhadap variabel terikat. Berdasarkan hasil analisis didapatkan nilai Adjusted R Square sebesar 0.554, yang berarti bahwa pengaruh evaluasi BSC melalui pendekatan ukuran umum dan unik dari masing-masing divisi dapat mempengaruhi besarnya alokasi kompensasi sebesar 55.4%, sedangkan sisanya

yaitu 44.6% dipengaruhi oleh variabel lain selain evaluasi BSC.

Tabel 4.7. Hasil Analisis Regresi

		Coefficients <sup>a</sup>						
		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
Model		B	Std. Error	Beta				Tolerance
1	(Constant)	1.402	.438		3.202	.004		
	Rc	.071	.025	.957	2.844	.009	.136	7.366
	Ru	.034	.013	.549	2.545	.017	.331	3.025
	Wc	-.059	.014	-.676	-4.102	.000	.567	1.764
	Wu	-.048	.012	-1.014	-3.837	.001	.220	4.544

a. Dependent Variable: Yrasio

Table diatas menunjukkan bahwa pengaruh ukuran umum dan unik masing-masing divisi terhadap alokasi nilai T hitung masing-masing sebesar 2.844 dengan signfiikasi sebesar 0.009 dimana nilai ini lebih kecil dari 0.05 ( $\alpha = 5\%$ ). Untuk Ru nilai T hitung sebesar 2.545 dengan signifikasi sebesar 0.017 lebih kecil dari 0.05 ( $\alpha = 5\%$ ). Untuk Wc nilai T hitung sebesar -4.102 dengan signifikasi sebesar 0.00 lebih kecil dari 0.05 ( $\alpha = 5\%$ ) dan untuk Wu nilai T hitung sebesar -3.837 dengan signifikasi sebesar 0.001 lebih kecil dari 0.05 ( $\alpha = 5\%$ ).

### KESIMPULAN

Hasil dari pengujian data dengan menggunakan profram SPSS 20 atas 30 partisipan yang dibagi dua lagi yaitu 15 partisipan yang mendapat pengetahuan BSC dan 15 partisipan yang tidak mendapat pengetahuan BSC ditemukan bahwa pendekatan *disaggregate /mechanically aggregate* berpengaruh dalam mengurangi bias pengukuran umum. Kesimpulan ini diambil berdasarkan *repeated measures ANOVA* yang menunjukkan bahwa penggunaan kedua metode pengukuran yaitu umum dan unik adalah sangat penting dalam menjelaskan hasil evaluasi kinerja secara keseluruhan. Dengan melihat nilai signfikansi effect antara interaksi divisi-common dan divisi-unique, maka dapat diketahui bahwa penggunaan pengukuran unik memiliki pengaruh yang lebih tinggi daripada pengukuran umum, hal ini diketahui dari nilai F hitung sebesar 9822.445 yang lebih besar jika dibandingkan dengan 937.626 untuk keseluruhan partisipan dan F hitung sebesar 20928.918 yang lebih besar dibandingkan 1935.214 untuk partisipan dengan pengetahuan. Serta untuk partisipan tanpa pengetahuan, pengukuran unik lebih

digunakan dibandingkan dengan pengukuran umum dibuktikan dengan F hitung sebesar 8173.753 untuk ukuran unik dan 809.554 untuk ukuran umum. Dengan melihat hasil tersebut maka hipotesis kedua dari penelitian dapat diterima atau dengan kata lain dapat dikatakan bahwa pengetahuan mengenai BSC dapat mengurangi bias pengukuran umum.

Pengetahuan tentang BSC dapat memberikan perbedaan yang signifikan pada evaluasi kinerja manajer baik dengan menggunakan pengukuran umum maupun pengukuran unik oleh partisipan penelitian. Hal ini terbukti dengan uji beda nilai evaluasi dengan pengukuran umum dan unik menghasilkan nilai signifikansi masing-masing sebesar 0.000 (sangat kecil) dimana nilai ini lebih kecil dari 0.05 ( $\alpha = 5\%$ ). Dari hasil ini maka dapat disimpulkan bahwa hipotesis kedua dari penelitian dapat diterima kebenarannya.

Dari hasil regresi pada table 4.14 diatas menunjukkan bahwa pengaruh ukuran umum dan unik masing-masing divisi terhadap alokasi kompensasi dapat dilihat dari nilai T hitung masing-masing sebesar 2.844 dengan signfiikasi sebesar 0.009 dimana nilai ini lebih kecil dari 0.05 ( $\alpha = 5\%$ ) untuk Rc. Untuk Ru nilai T hitung sebesar 2.545 dengan signifikasi sebesar 0.017 lebih kecil dari 0.05 ( $\alpha = 5\%$ ). Untuk Wc nilai T hitung sebesar -4.102 dengan signifikasi sebesar 0.00 lebih kecil dari 0.05 ( $\alpha = 5\%$ ) dan untuk Wu nilai T hitung sebesar -3.837 dengan signifikasi sebesar 0.001 lebih kecil dari 0.05 ( $\alpha = 5\%$ ). Sehingga dapat disimpulkan baik pengukuran umum maupun pengukuran unik berpengaruh secara signifikan terhadap alokasi bonus. Dari hasil ini maka dapat disimpulkan bahwa hipotesis ketiga dari penelitian ini dapat diterima kebenarannya.

### Keterbatasan dan Saran untuk Penelitian Selanjutnya

Keterbatasan dalam penelitian ini adalah sebagai berikut:

1. Penelitian ini tidak mengukur seberapa besar peningkatan pengetahuan partisipan mengenai BSC sebelum dan sesudah diberikan pengetahuan.
2. Penelitian ini belum menggunakan pendekatan-pendekatan lain seperti

pendekatan akuntabilitas yang digunakan oleh Libby, Salterio & Webb (2004) dan pendekatan pelatihan kepada partisipan yang digunakan oleh Dilla & Steinbart (2005) yang dapat mengurangi bias pengukuran umum dalam penilaian kinerja menggunakan BSC.

3. Partisipan dalam eksperimen belum berpengalaman dalam menggunakan BSC.

Saran yang dapat diberikan terkait penelitian ini adalah sebagai berikut:

1. Dalam penelitian selanjutnya, peneliti diharapkan untuk mengadakan evaluasi yang berisi tentang pertanyaan-pertanyaan terkait teori dan penerapan BSC kepada partisipan sebagai alat untuk mengukur peningkatan pengetahuan sebelum dan sesudah diberikan review materi BSC.
2. Penelitian selanjutnya disarankan untuk menggunakan kombinasi pendekatan yang lebih banyak sebagaimana yang telah disebutkan dalam keterbatasan untuk mengurangi bias.
3. Penelitian selanjutnya disarankan untuk memberikan pelatihan mengenai bagaimana membuat dan menggunakan BSC.

#### DAFTAR PUSTAKA

- Anderson, J. R. (1990). *Cognitive Psychology and its Implications: Third Edition*. New York: Freeman.
- Banker, D. R., Chang, H., & Pizzini, M. (2004). The Balanced Scorecard: Judgmental Effects of Performance Measures Linked to Strategy. *The Accounting Review*, 79(1), 1-23.
- Bazerman, M. H., & Moore, D. A. (2009). *Judgment in Managerial Decision Making* (7th.ed.). New Jersey: John Wiley & Sons.
- Bone, H. & Sholihin, M. (2012). Pengaruh Perspektif Dan Jenis Ukuran Dalam Balanced Scorecard Terhadap Evaluasi Kinerja. *Thesis, Universitas Gadjah Mada*.
- Bonner, S. E. (1990). Experience Effects in Auditing: The Role of Task-Specific Knowledge. *The Accounting Review*, 65 (1), 72-92
- Bonner, S. E., & Lewis, B. L. (1990). Determinants of Auditor Expertise. *Journal of Accounting and Auditing*, 28, 1-20
- Bowman, E. H. (1963). Consistency and Optimality in Managerial Decision Making. *Management Science*, 9 (1), 31-321.
- Dilla, W. N., & Steinbart, P. J. (2005a). Relative Weighting of Common and Unique Measures by Knowledgeable Decision Makers. *Behavioral Research in Accounting*, 17 (1), 43-53.
- Einhorn, H. J. (1974, Oktober). Expert Judgment: Some Necessary Conditions and an Example. *Journal of Applied Psychology*, 71-562.
- Gagne, M. L., Hollister, J., & Tully, G. J. (2006). Using the Balanced Scorecard: Both Common and Unique Measures are Informative. *Journal of Applied Business Research*, 22(1), 147-160.
- Gibbs, M. J, Merchant, K. A., Wim A Van der Stede, & Vargus, M. E. (2005). The Benefits of Evaluating Performance Subjectively. *Performance Improvement*, 44 (5), 26-32
- Grevinga, K. H. M. (2013). Common Measure Bias in the Balanced Scorecard: an Experiment with Undergraduate Students. *Master's Degree Thesis, University of Twente, Netherlands*.
- Holmstrom, B., & Milgrom, P. (1991). *Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design*. *Jurnal of Law, Economics, & Organization*, 7, 24-52.
- Humphreys, K. A., & Trotman, K. T. (2011). *The Balanced Scorecard: The Effect of Stratey Information on Perfomance*

- Evaluation Judgments. Journal of Management Accounting Research*, 23, 81-98.
- Indriantoro, N., & Supomo, B. (1999, Oktober). Metode Penelitian Bisnis untuk Akuntansi dan Manajemen, BPFE-Yogyakarta, Edisi Pertama.
- Ittner, C. D., Larcker, D. F., & Meyer, M. W. (2003). Subjectivity and the Weighting of Performance Measures: Evidence from a Balanced Scorecard. *The Accounting Review*, 78 (3), 725-758.
- Johnson, E. N., Reckers, P. M. J., & Barlett, G. D. (2013). Influence Timeline and Perceived Strategy Effectiveness On Balance Scorecard Performance Evaluation Judgements. *Journal of Management Accounting Research*.
- Kaplan, R. S., & Norton, D. P. (1992). The Balanced Scorecard - Measures that Drive Performance. *Harvard Business Review*, 70 (1), 71-79.
- Kaplan, R. S., & Norton, D. P. (1996). Linking the Balanced Scorecard to Strategy. *California Management Review*, 39 (1), 53-79.
- Kaplan, R. S., & Norton, D. P. (2001). The Strategy - Focused Organization: How Balanced Scorecard Companies Thrive in the New Business Environment. *Harvard Business School Press*.
- Kaplan, R. S., & Norton, D. P. (2001a). Transforming the Balanced Scorecard from Performance Measurement to Strategic Management (Part I). *Accounting Horizons*, 15 (1), 87-104.
- Kaplan, R. S., & Norton, D.P. (2001b). Transforming the Balanced Scorecard from Performance Measurement to Strategic Management (Part II). *Accounting Horizons*, 15 (2), 147-160.
- Kaskey, V. L. (2008). The Balanced Scorecard: A Comparative Study of Accounting Education and Experience on Common Measure Bias and Trust in a Balanced Scorecard. *PhD Dissertation. School of Business and Technology, Capella University, USA*.
- Kennedy, J. (1995). Debiasing the Curse of Knowledge in Audit Judgment. *The Accounting Review*, 70 (2), 249-273.
- Libby, T., Salterio, S., & Webb, A. (2002). The Balanced Scorecard: The Effect of Assurance and Process Accountability on Managerial Judgment.
- Liedtka, S. L., Church, B. K., & Ray, M. R. (2008). Performance Variability, Ambiguity Intolerance, and Balanced Scorecard-Based Performance Assessments. *Behavioral Research in Accounting*, 20 (2), 73-88.
- Lipe, M. G., & Salterio, S. E. (2000). The Balanced Scorecard: Judgmental Effects of Common and Unique Performance Measures. *The Accounting Review*, 75 (3), 283-298.
- Malina, M. A., & F. H. Selto. (2001). Communicating and controlling strategy. An empirical study for the effectiveness of the balanced scorecard. *Journal Management Accounting Research* 13, 47-90.
- Nelson, R. R. (1987). The Role of Knowledge in R&D Efficiency. *Oxford University Press*, 97 (3), 453-470.
- Niven, P. R. (2006). *Balanced Scorecard Step-by-Step Maximizing Performance And Maintaining Results (2th. Ed.)* New Jersey: John Wiley & Sons.
- Payne, J., J. Bettman & E. Johnson. (1993). *The Adaptive Decision Maker*. Cambridge, U. K.: Cambridge University Press.
- Roberts, M. L., Albright, T. L., & Hibbets, A. R. (2004). Debiasing Balanced Scorecard Evaluations. *Behavioral Research in Accounting*, 16 (1), 75-88.
- Serfas, S. (2011). Cognitive Biases in the Capital Investment Context: Theoretical Consideration and Empirical Experiments on Violations of Normative Rationality. *Dissertation, Chemnitz University of Technology, Germany*. Retrieved May 17, 2014, from <http://books.google.co.id/books?id=i7OJWje1JgQC&printsec=frontcover#v=onepage&q&f=false>.
- Slovic, P., & MacPhillamy, D. (1974). Dimensional Commensurability and Cue Utilization in Comparative Judgment. *Organizational Behavior*

*and Human Performance*, 11 (2), 172-194.

Tversky, A., & Kahneman, D. (1973). Availability: A Heuristic for Judging Frequency and Probability. *Cognitive Psychology*, 5, 207-232.

Tversky, A., & Kahneman, D. (1974). Judgment Under Uncertainty: Heuristics and Biases. *Science*, 185, 1124-1131.