

# Korelasi Antara Validitas Pada Evaluasi Yang Digunakan Dalam Menilai Hasil Belajar Siswa Dengan Hasil Kegiatan MGMP Matematika Di Kabupaten Pidie

Miftahuddin\* dan Fithriana AR\*

## Abstrak

Validitas dapat dijadikan alternatif dalam mengevaluasi hasil belajar siswa dalam proses pembelajaran matematika di sekolah. Terdapat korelasi searah antara hasil *pre-post test* pada kegiatan MGMP (Musyawarah Guru Mata Pelajaran) matematika tingkat SMP/MTs tahun 2006/2007 di Kabupaten Pidie dengan Modus ( $M_o$ ) pada evaluasi belajar siswa kelas III. Pengujian validitas dilakukan dengan menggunakan metode korelasi *product moment*. Korelasi antara nilai ujian sumatif dan nilai ulangan umum mata pelajaran matematika memiliki tingkat validitas yang tinggi secara kuantitatif dengan taraf signifikan  $\alpha = 1\%$ .

**Kata Kunci :** *Validitas, evaluasi belajar, korelasi Product Moment,  $M_o$ , kuantitatif.*

## 1. Pendahuluan

Untuk mempersiapkan sumber daya manusia yang berkualitas hanya dapat diperoleh dengan meningkatkan mutu pendidikan. Dengan kata lain pendidikan merupakan upaya paling urgen untuk mempersiapkan anak bangsa menghadapi persaingan global. Menyadari hal ini, berbagai upaya kebijakan pendidikan dilakukan oleh pemerintah. Salah satunya adalah melalui kegiatan MGMP dan ujian lokal yang dilakukan guru-guru mata pelajaran, sehingga dapat diketahui gambaran kualitas pendidikan yang dijalankan pada suatu sekolah.

Belajar merupakan suatu proses dimana keberhasilannya sangat ditentukan oleh faktor internal dan eksternal. Faktor internal merupakan faktor yang ada dalam diri siswa seperti intelegensi, minat dan bakat serta motivasi. Faktor eksternal mencakup faktor guru, lingkungan keluarga, masyarakat, dan pemerintah. Menjadi pemikiran bahwa sejauh faktor internal dan eksternal tersebut dapat diatasi atau diantisipasi, sehingga siswa dapat mencapai prestasi yang baik. Guru sebagai salah satu faktor penting yang ikut menentukan prestasi belajar siswa. Jika menguasai materi pelajaran dengan baik, menggunakan bermacam media belajar dan menerapkan variasi metode dalam mengajar dapat menyebabkan siswa menjadi senang untuk belajar. Guru juga memberikan penilaian atau evaluasi belajar kepada siswa. Dalam hal ini guru seharusnya membuat alat ukur yang valid. Tetapi dalam pelaksanaannya masih terdapat kelemahan dalam proses penilaian siswa di sekolah. Faktor penyebabnya antara lain adalah evaluasi yang diberikan guru kurang tepat (kurang valid) atau kurang memenuhi syarat sebagai suatu alat ukur. Oleh karena itu, timbul suatu permasalahan yaitu apakah bentuk evaluasi yang telah dibuat memenuhi syarat atau tidak. Dalam hal ini terdapat keterbatasan menguji validitas alat ukur maka yang dilakukan adalah mengukur kesahihan butir (validitas item).

---

\* Staf Pengajar pada Jurusan Matematika, FMIPA Universitas Syiah Kuala, Nangroe Aceh Darussalam  
Email: [miftah@unsyiah.ac.id](mailto:miftah@unsyiah.ac.id), [fithriana.ar@unsyiah.ac.id](mailto:fithriana.ar@unsyiah.ac.id)

Korelasi juga disebut koefisien korelasi, adalah nilai yang menunjukkan kekuatan dan arah hubungan linier antara dua peubah acak. Koefisien antara evaluasi dan kriteria itu ( $r_{xy}$ ) inilah yang digunakan untuk menyatakan tinggi-rendahnya validitas suatu alat ukur. Pengujian dilakukan dengan menggunakan korelasi *product moment* dari Pearson. Penelitian ini bertujuan untuk mengetahui valid tidaknya evaluasi dari ujian lokal yang digunakan guru matematika dalam mengevaluasi belajar siswa dan hasil kegiatan MGMP matematika SMP/MTs di Kabupaten Pidie tahun pelajaran 2006/2007.

## 2. Pembahasan

Untuk mempelajari teori validitas dapat diketahui dari kata ‘*valid*’ yang artinya tepat atau sah. Suatu alat ukur dapat digunakan sebagai alat ukur yang valid apabila alat pengukur tersebut dapat mengukur apa yang hendak diukur secara tepat dan teliti. Sebuah evaluasi disebut valid apabila evaluasi itu dapat mengukur apa yang hendak diukur (lihat [2]). Walaupun telah dikemukakan bahwa validitas dapat menunjukkan kualitas tinggi sampai rendah, namun hal ini masih sangat umum. Para ahli menggolongkan validitas ke dalam beberapa jenis. Jenis-jenis validitas (lihat [8]) diuraikan berikut ini.

### *Content Validity*

Suatu evaluasi dikatakan memiliki *content validity* apabila evaluasi tersebut dapat mewakili seluruh materi kurikuler yang telah dilaksanakan. Validitas ini sering disebut validitas isi. Validitas ini menunjuk pada evaluasi yang bersampel representatif dengan materi pelajaran yang telah diberikan.

### *Construct Validity*

Untuk menentukan adanya *construct validity*, suatu evaluasi dikorelasikan dengan konsep atau teori, dimana item-item soal tersebut harus sesuai dengan ciri-ciri konsep yang akan dievaluasi. Dengan kata lain, soal-soal tersebut disesuaikan dengan indikasi-indikasi lain dari ciri-ciri variabel yang akan diukur. Validitas konstruksi juga disebut dengan validitas logis sebagai kualitas anak. Kemampuan berpikir anak secara logis. Kemampuan berpikir secara logis dapat diketahui dengan hasil tes dengan skor tinggi. Peralnya inteligensi, evaluasi bakat dapat mengukur aspek-aspek yang menonjol pada bakat siswa.

### *Predictif Validity*

Validitas tes yang digunakan untuk meramalkan kemampuan siswa pada masa yang akan datang disebut dengan validitas prediksi. Validitas prediksi menunjukkan kesesuaian antara skor evaluasi dengan variabel yang akan diukur untuk keadaan pada masa yang akan datang. Skor evaluasi dapat digunakan sebagai indikator untuk memprediksikan keberhasilan belajar siswa di masa mendatang dalam waktu yang relatif lama. *Predictif validity* sangat penting untuk melihat dan mengklasifikasi individu-individu kedalam jenjang-jenjang pendidikan yang sesuai dengan bakat, minat dan kemampuan. Suatu evaluasi dikatakan “*prediktif validity*” jika hasil korelasi dapat meramalkan dengan tepat keberhasilan seseorang di masa yang akan datang dalam lapangan tertentu. Tepat tidaknya ramalan itu dapat dilihat dari korelasi koefisien antara hasil evaluasi dengan alat pengukur lain kelak di masa mendatang.

*Concurrent Validity*

Suatu evaluasi mempunyai korelasi yang tinggi dengan hasil pengukuran lain terhadap bidang yang sama dan dalam waktu yang sama pula, maka evaluasi itu dikatakan memiliki *concurrent validity*. Perlu diingat bahwa *prediktif validity* dan *concurrent validity* mempunyai kesamaan, dimana antara skor dengan alat pengukur lain lebih menunjukkan hubungan atau korelasi yang lebih tinggi, namun pada prinsipnya *prediktif validity* mengarah pada hubungan yang terjadi pada masa yang akan datang, sedangkan *concurrent validity* mengarah pada hubungan yang terjadi pada saat ini atau pada keadaan sekarang. Kedua validitas ini disebut juga dengan dengan validitas kriteria.

Skor evaluasi yang digunakan disebut prediktor, sedangkan skor evaluasi yang digunakan untuk memprediksi keberhasilan atau perubahan yang terjadi disebut dengan kriterium. Jadi validitas kriteria diperoleh dengan cara membandingkan antara prediktor dengan kriterium. Untuk memperoleh perbandingan tersebut dilakukan dengan koefisien korelasi yang besarnya antara -1 s/d +1. Tinggi rendahnya koefisien korelasi dapat ditentukan dengan angka koefisien korelasi berkisar antara beberapa tingkat yaitu: (1). 0.00 – < 0.20 : korelasi sangat rendah; (2). 0.20 – < 0.40: korelasi rendah; (3). 0.40 – < 0.70 : korelasi cukup; (4). 0.70 – < 0.90 : korelasi tinggi; (5). 0.90 – 1.00 : korelasi sangat tinggi (lihat [9]).

## 2.1 Tes sebagai Alat Evaluasi

Tes yang disebut juga alat ukur merupakan salah satu cara untuk mengadakan evaluasi. Evaluasi juga dapat dilakukan dengan cara non-tes seperti wawancara dan pengamatan. Tes adalah suatu cara untuk mendapatkan penilaian yang berbentuk suatu tugas atau serangkaian tugas harus dikerjakan oleh anak atau sekelompok anak sehingga menghasilkan suatu nilai tentang tingkah laku atau prestasi anak tersebut, yang oleh anak-anak lain atau dengan nilai standar yang ditetapkan (lihat [7]).

Evaluasi adalah suatu proses berkelanjutan tentang pengumpulan informasi dan penafsiran informasi untuk menilai keputusan-keputusan yang dibuat dalam merancang suatu sistem belajar. Rumusan itu mempunyai tiga implikasi, yaitu :

- (1) Evaluasi adalah suatu proses yang terus menerus, bukan hanya saja pada akhir pengajaran, tetapi dimulai sebelum dilaksanakannya pengajaran sampai dengan berakhirnya pengajaran.
- (2) Proses evaluasi senantiasa diarahkan ke tujuan tertentu, yakni untuk mendapatkan jawaban-jawaban tentang bagaimana memperbaiki pengajaran.
- (3) Evaluasi menuntut penggunaan alat-alat ukur yang akurat dan bermakna untuk mengumpulkan informasi yang dibutuhkan untuk membuat keputusan. Dengan demikian, evaluasi merupakan proses yang berhubungan dengan pengumpulan informasi yang memungkinkan untuk menentukan tingkat kemajuan pengajaran (lihat [3]).

Evaluasi pada umumnya mengandung fungsi dan tujuan yaitu :

1. Untuk menentukan angka kemajuan atau hasil belajar para siswa. Angka-angka yang diperoleh dicantumkan sebagai laporan kepada orang tua, untuk kenaikan kelas, dan penentuan kelulusan para siswa.
2. Untuk menempatkan para siswa ke dalam situasi belajar mengajar yang tepat dan serasi dengan tingkat kemampuan, minat, dan berbagai karakteristik yang dimiliki oleh setiap siswa.
3. Untuk mengenal latar belakang siswa (psikologis, fisik, dan lingkungan), yang berguna untuk menentukan sebab-sebab kesulitan belajar para siswa. Informasi yang diperoleh dapat digunakan untuk memberikan bimbingan dan penyuluhan pendidikan untuk mengatasi kesulitan-kesulitan yang mereka hadapi.

4. Sebagai umpan balik bagi guru yang pada gilirannya dapat digunakan untuk memperbaiki proses belajar mengajar.

Fungsi pertama umumnya banyak mendapat perhatian dan pelaksanaan pengajaran sehari-hari. Padahal fungsi-fungsi lainnya tidak kalah pentingnya, bahkan memegang peranan yang cukup menentukan terhadap keberhasilan pendidikan siswa dalam jangka waktu yang lama (lihat [3]). Untuk mengukur prestasi belajar siswa, dibutuhkan suatu alat ukur yang akurat yang dapat diandalkan. Jika tidak maka informasi yang diperoleh tidak dapat dipercaya dan mungkin tidak memberikan gambaran yang sebenarnya tentang hasil belajar siswa. Itu sebabnya penyusunan suatu instrumen harus memenuhi persyaratan tertentu seperti validitas (lihat [3]).

## 2.2 Validitas Tes

Validitas tes adalah ketetapan tes terhadap konsep yang dinilai sehingga betul-betul menilai atau mengukur apa yang seharusnya diukur atau dinilai. Bila hal demikian terwujud maka tes tersebut dikatakan valid atau tepat. Sebuah tes disebut valid apabila tes ini dapat tepat mengukur apa yang hendak diukur. Istilah valid sangat sukar dicari ganti, ada istilah baru yang diperkenalkan, yaitu sah, sehingga belum dapat mencakup semua arti yang tersirat dalam kata "Valid", dan kata "Tepat" kadang-kadang digunakan dalam konteks yang lain, akan tetapi tambahan kata "tepat" dalam menerangkan kata "valid" dapat memperjelas apa yang dimaksud. Berdasarkan penjelasan dan tanggapan di atas, maka dapat di ambil kesimpulan bahwa validitas tes adalah ketepatan alat ukur dengan objek yang dinilai atau diukur juga dengan tujuan penilaian dan pengukuran (lihat [2]).

## 2.3 Uji Validitas

Validitas didefinisikan sebagai seberapa cermat suatu alat tes melakukan fungsi ukurnya atau bisa juga diartikan sejauh mana suatu alat ukur dapat mengukur apa yang ingin diukur (lihat [10]). Pengukuran ini dapat dilakukan dengan menggunakan teknik korelasi *Product Moment* dari Pearson. Teknik ini bertujuan untuk menguji apakah tiap item atau butir pernyataan benar-benar mampu mengungkap faktor yang akan diukur atau konsistensi internal tiap item alat ukur dalam mengukur suatu faktor. Nilai korelasi yang diperoleh lalu dibandingkan dengan tabel nilai korelasi ( $r$ ) *Product Moment* untuk mengetahui apakah nilai korelasi yang diperoleh signifikan atau tidak. Jika indeks nilai yang diperoleh dari perhitungan tersebut memiliki nilai yang lebih besar dari nilai tabel korelasi maka item itu dinyatakan valid demikian juga sebaliknya.

## 2.4 Teknik Korelasi *Product Moment*

*Product moment corellation* adalah salah satu teknik untuk mencari korelasi antar dua variabel yang kerap kali digunakan. Teknik korelasi ini dikembangkan oleh Karl Pearson, yang karenanya sering dikenal dengan istilah Teknik Korelasi Pearson. Korelasi yang terjadi antara dua variabel dapat berupa korelasi positif, korelasi negatif, tidak ada korelasi, ataupun korelasi sempurna. Disebut *product moment corellation* karena koefisien korelasinya diperoleh dengan cara mencari hasil perkalian dari momen-momen variabel yang dikorelasikan (*product of the moment*) (lihat [4]). Jika sampel yang kita teliti merupakan sampel besar ( $N \geq 30$ ), maka cara mencari atau menghitung angka indeks korelasi " $r$ " *product moment* dilakukan dengan menggunakan alat bantu berupa peta korelasi atau *scatter diagram* (lihat [9]). Rumus yang digunakan :

$$r_{xy} = \frac{\frac{\sum x' y'}{N} - (C_x')(C_y')}{(SD_x')(SD_y')} \quad (1)$$

dimana:

$r_{xy}$  = Korelasi "r" *product moment* atau tingkat validitas

$\sum x' y'$  = Jumlah hasil perkalian silang (*product of the moment*) antara frekuensi ( $f$ ) dengan  $x'$  dan  $y'$ .

$C_x'$  = Nilai korelasi pada nilai ulangan umum ( $Y$ )

$C_y'$  = Nilai korelasi pada nilai ujian sumatif ( $X$ )

$SD_x'$  = Standar deviasi skor  $X$  dalam arti tiap skor sebagai 1 unit (dimana  $i = 1$ )

$SD_y'$  = Standar deviasi skor  $Y$  dalam arti tiap skor sebagai 1 unit (dimana  $i = 1$ )

$N$  = Jumlah sampel

Terdapat beberapa langkah prosedur yang perlu ditempuh, yaitu

Langkah (1) Peta korelasi (*scatter diagram*);

(2) Mencari  $C_x'$  dengan rumus :  $C_x' = \frac{\sum fx'}{N}$  (2)

dimana:

$\sum fx'$  = Jumlah hasil perkalian antara frekuensi masing-masing skor dengan  $x'$

(3) Mencari  $C_y'$  dengan rumus :  $C_y' = \frac{\sum fy'}{N}$  (3)

dimana:

$\sum fy'$  = Jumlah hasil perkalian antara frekuensi masing-masing skor dengan  $y'$ .

(4) Mencari  $SD_x'$  dengan rumus :  $SD_x' = i \sqrt{\frac{\sum fx'^2}{N} - \left(\frac{\sum fx'}{N}\right)^2}$  (4)

dimana:

$\sum fx'^2$  = Jumlah hasil perkalian antara frekuensi masing-masing skor dengan  $x'^2$ .

(5) Mencari  $SD_y'$  dengan rumus :  $SD_y' = i \sqrt{\frac{\sum fy'^2}{N} - \left(\frac{\sum fy'}{N}\right)^2}$  (5)

dimana:

$\sum fy'^2$  = Jumlah hasil perkalian antara frekuensi masing-masing skor dengan  $y'^2$ .

(6) Mencari  $r_{xy}$  dengan rumus (1).

(7) Memberi interpretasi terhadap  $r_{xy}$ .

Terlebih dahulu dirumuskan hipotesis  $H_o$  dan  $H_a$ .

## 2.5 Tabel Tabulasi Silang

Tabel tabulasi silang (*cross tabulation table*), merupakan cara deskriptif sederhana untuk melihat apakah ada hubungan antara dua buah variabel. Dengan tabel silang akan lebih mudah mengilustrasikan hubungan antara dua variabel (lihat [5]). *Crosstab analysis* merupakan alat uji statistika untuk data yang bersifat kualitatif atau non parametrik (lihat [11]).

Dalam hal ini yang menjadi objek penelitian adalah butir-butir *test* matematika dan yang menjadi subyeknya adalah siswa SMP N 4 Bandar Baru Pidie. Mengingat siswa SMP N 4 Bandar Baru terdiri dari kelas I, II, III, maka hanya diambil sebagian saja yaitu kelas III saja sebagai populasinya. Jumlah populasi dalam penelitian ini adalah 160 siswa yang tersebar dalam 4 kelas. Data yang digunakan dalam penulisan ini merupakan data sekunder dari Nilai Ulangan Umum (NUM) dan Nilai Ujian Sumatif (NUS) mata pelajaran matematika siswa SMPN 4 Bandar Baru Pidie. Selanjutnya digunakan hasil evaluasi dari kegiatan MGMP pada guru-guru matematika SMP/MTs yang dipilih secara acak di Kabupaten Pidie tahun pelajaran 2006/2007.

## 3. Aplikasi

Adapun data-data yang diperoleh dalam penelitian ini adalah

- (1). Nilai mata pelajaran matematika NUM dan NUS semester ganjil kelas III di SMPN 4 Bandar Baru Pidie tahun pelajaran 2006/2007.
- (2). Soal ujian matematika kelas III semester ganjil SMPN 4 Bandar Baru Pidie berbentuk soal objektif sebanyak 25 butir dan soal esai 5 butir.
- (3). Adapun NUM dan NUS mata pelajaran matematika di SMPN 4 Bandar Baru Pidie dapat dilihat pada Sumber : SMP Negeri 4 Bandar Baru-Pidie (2006/2007).
- (4). Hasil instrumen (data pendukung) dapat dilihat pada Tabel 6.
- (5). Hasil kegiatan (*initial baseline* dan *follow up survey*) MGMP Matematika tahun 2006/2007 di Kabupaten Pidie, dari total guru matematika SMP sebanyak 154 orang dan guru matematika MTs sebanyak 120 orang.

Pengolahan data dilakukan untuk mendapatkan tingkat validitas dari soal-soal yang dibuat guru mata pelajaran matematika di SMPN 4 Bandar Baru-Pidie dan soal-soal yang diberikan pada guru-guru matematika dalam kegiatan MGMP tahun pelajaran 2006/2007 serta korelasi keduanya. Perhitungan tingkat validitas dilakukan dengan cara membandingkan data NUM ( $Y$ ) dengan NUS ( $X$ ). NUM adalah nilai dari ujian akhir semester, sedangkan NUS adalah nilai dari ujian harian. Hasil dari perhitungan ini berguna dalam rangka kegiatan diagnostik dan penempatan siswa. Diagnostik berfungsi sebagai pemberian bimbingan kepada siswa yang menghadapi kesulitan belajar. Penempatan siswa berhubungan dengan penentuan diterima atau tidaknya siswa pada sekolah tertentu dan di kelas yang sesuai dengan informasi tentang siswa yang bersangkutan.

### Menentukan Koefisien Korelasi atau Tingkat Validitas

Dari data awal, pertama dicari rata-rata dari NUS dan rata-rata NUM. Cara memperoleh rata-rata dari NUS dan NUM sebagai berikut :

$$\bar{X}_{A(NUS)1} = \frac{\sum_{i=1}^{40} x_{i1}}{40} = 65.8,$$

$$\bar{X}_{A(NUS)2} = \frac{\sum_{i=1}^{40} x_{i2}}{40} = 67.9,$$

$$\bar{X}_{A(NUS)3} = \frac{\sum_{i=1}^{40} x_{i3}}{40} = 71.5,$$

$$\bar{X}_{A(NUS)4} = \frac{\sum_{i=1}^{40} x_{i4}}{40} = 77.3,$$

$$\bar{X}_{A(NUS)5} = \frac{\sum_{i=1}^{40} x_{i5}}{40} = 64.9, \text{ jumlah rata-rata NUS yang ke-5 pada kelas IIIA.}$$

Hasil perhitungan seluruhnya disajikan pada [6] Lampiran A. Dari nilai rata-rata ujian sumatif dapat dikelompokkan lagi menjadi 1 yang terbaik, 2 yang terbaik, 3 yang terbaik dan 4 yang terbaik dari tiap-tiap kelas.

$$\bar{X}_{A(UNS)1} = \frac{\sum_{i=1}^{40} x_{1i}}{40} = 77.3, \text{ jumlah rata-rata NUS 1 yang terbaik pada kelas III A.}$$

Sedangkan

$$\bar{Y}_{A(NUM)} = \frac{\sum_{j=1}^{40} y_j}{40} = 75.95, \text{ jumlah rata-rata NUM kelas III A.}$$

NUS terbaik terdapat pada kolom 4 dan NUS terbaik kedua ada di kolom ketiga.

$$68 \approx \frac{65 + 70}{2}, 73 \approx \frac{80 + 65}{2}, \dots$$

NUS 2 terbaik diletakkan pada kolom kedua kelas III A. Dari hasil di atas dicari nilai rata-rata NUS 2 yang terbaik, yaitu

$$\bar{X}_{A(NUS)2} = \frac{\sum_{i=1}^{40} x_{2i}}{40} = 74.7, \text{ yang merupakan jumlah rata-rata NUS 2 yang terbaik pada}$$

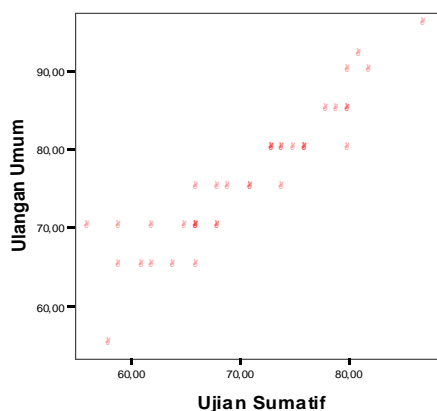
kelas III A. Hasil lengkapnya dapat dilihat pada tabel berikut.

Tabel 1. Rata-rata NUS dan NUM Mata Pelajaran Matematika di SMPN 4 Bandar Baru

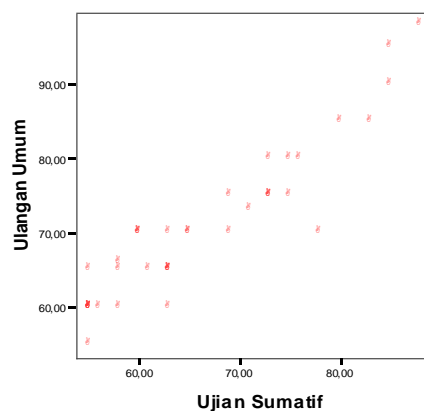
No	Kelas III A					...	Kelas III D				
	NUS				NUM		NUS				NUM
	1	2	3	4			1	2	3	4	
1	65	68	62	61	65	...	60	65	60	60	50
2	80	73	72	66	75	...	80	75	77	75	65
3	80	80	77	73	80	...	60	50	53	50	60
4	80	80	77	74	80	...	60	55	60	60	50
5	70	60	60	59	65	...	70	75	73	70	65

⋮	⋮	⋮	⋮	⋮	⋮	...	⋮	⋮	⋮	⋮	⋮
36	80	83	85	87	96	...	70	70	67	65	65
37	90	88	82	82	90	...	70	60	60	60	50
38	65	58	59	56	70	...	70	65	63	60	55
39	75	73	68	66	70	...	80	75	70	70	65
40	75	65	63	59	70	...	70	65	60	60	60
$\mu$	77.3	74,7	72.2	70.7	75.95	...	70	67.8	65	63.5	59.225

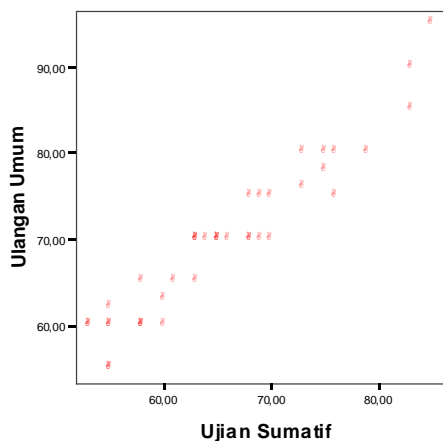
Catatan : Hasil perhitungan seluruhnya ditunjukkan pada [6] Lampiran C.



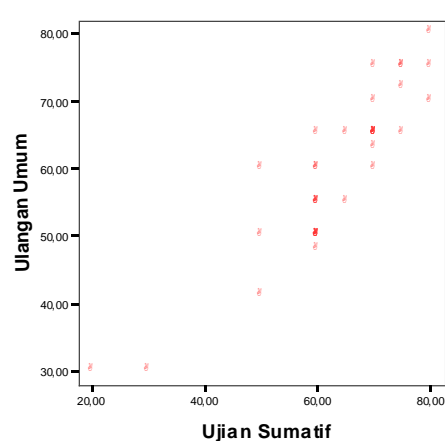
(a)



(b)



(c)



(d)

Gambar 1. Peta Korelasi dari empat NUS yang Terbaik dan NUM  
(a) Kelas III A, (b) Kelas III B, (c) Kelas III C, (d) Kelas III D (berdasarkan nilai yang diperoleh siswa)

Gambar di atas menunjukkan hubungan yang searah antara ulangan umum dengan ujian sumatif. Semakin tinggi NUS maka semakin tinggi NUM. Hal ini dilihat dari nilai korelasi ke empat kelas



III. Selanjutnya dapat dicari nilai korelasi dari keempat kelas III, dari satu NUS terbaik sampai empat NUS yang terbaik, dan kemudian akan dikorelasikan dengan NUM.

Tabel 2. Korelasi NUS dan NUM Matematika di SMP N 4 Bandar Baru-Pidie

Kelas	Korelasi Karl - Pearson				$r_{tabel}$
	1 dan NUM	2 dan NUM	3 dan NUM	4 dan NUM	
A	0.769	0.826	0.920	0.924	0.403
B	0.715	0.867	0.925	0.930	0.403
C	0.786	0.917	0.923	0.957	0.403
D	0.737	0.760	0.889	0.888	0.403

Dari Tabel 2 dapat dilihat untuk kelas III A, III B, III C, nilai korelasi yang paling bagus adalah nilai dari korelasi antara 4 NUS yang terbaik dengan NUM yaitu masing-masing 0.924, 0.93, 0.957. Kelas III D nilai korelasi yang paling bagus adalah nilai dari korelasi antara 3 NUS yang terbaik dengan NUM 0.889, tetapi karena ada 2 siswa yang tidak mengikuti ujian sumatif sebanyak 4 kali, jadi nilai korelasi yang diambil adalah nilai dari korelasi antara 4 NUS yang terbaik dengan NUM 0.888. Untuk memberikan interpretasi terhadap  $r_{xy}$ , dirumuskan hipotesis berikut.

$H_0$  : Tidak ada korelasi positif signifikan antara variabel  $X$  dan variabel  $Y$

$H_a$  : Ada korelasi positif yang signifikan antara variabel  $X$  dan variabel  $Y$

Dengan  $N$  sebesar 40 diperoleh  $r_{tabel}$  pada taraf signifikan 1% sebesar 0.403. Kelas III A mempunyai nilai  $r_{xy(\text{hitung})} = 0.924 > r_{tabel} = 0.403$ . Karena  $r_{xy(\text{hitung})} = 0.924 > r_{tabel} = 0.403$ , maka  $H_0$  ditolak. Kelas III B mempunyai  $r_{xy(\text{hitung})} = 0.93 > r_{tabel} = 0.403$ , sehingga  $H_0$  ditolak. Kelas III C mempunyai nilai  $r_{xy(\text{hitung})} = 0.957 > r_{tabel} = 0.403$  maka  $H_0$  ditolak. Kelas III D mempunyai  $r_{xy(\text{hitung})} = 0.888 > r_{tabel} = 0.403$ , sehingga  $H_0$  ditolak. Berarti terdapat korelasi positif yang signifikan antara variabel  $X$  dan variabel  $Y$ . Kesimpulannya, tinggi rendahnya NUS menentukan tinggi rendahnya NUM mata pelajaran matematika di SMPN 4 Bandar Baru-Pidie tahun pelajaran 2006/2007. Atau dengan kata lain, evaluasi yang digunakan guru matematika dalam menilai hasil belajar siswa sudah valid secara kuantitatif.

Tabel 6 data *crosstab* memberikan penjelasan bahwa kelas D korelasinya lebih kecil dari pada kelas A, B dan C. Hal ini disebabkan oleh kelas A mendapat dukungan dari orang tua atau masyarakat selama belajar matematika di rumah sebesar 37.5%, kelas B sebesar 75%, kelas C sebesar 62.5%, sedangkan kelas D sebesar 25%.

Dari uraian di atas dapat disimpulkan bahwa validitas yang digunakan oleh guru di SMPN 4 Bandar Baru-Pidie adalah *content validity* dan *construct validity* dimana tugas yang diberikan guru berasal dari buku panduan. *Concurrent validity* juga digunakan, dimana pengukuran dilakukan pada bidang yang sama (mata pelajaran matematika) dan dalam waktu yang sama (dalam satu semester). Serta *prediktif validity* dimana dapat digunakan untuk meramal kemampuan siswa pada masa yang akan datang. Tabel berikut memperlihatkan hasil perhitungan korelasi pada kelas A, B, C dan D.

Tabel 3. Modus, Jumlah dan Persentase NUS dan NUM di Kelas III A, B, C, dan D

	Kelas	Mean	Stdv	Modus	$\Sigma$	%
NUS 1	A	65.775	9.963	65	9	22.50
NUS 2		67.790	9.758	70	12	30.00
NUS 3		71.500	9.819	70	12	30.00
NUS 4		77.250	7.675	80	12	30.00

NUS 5		64.900	12.341	65	25	20.00
NUM	A	75.950	8.688	70 & 80		
NUS 1	B	66.250	12.897	60	14	35.00
NUS 2		66.000	10.931	60	15	37.50
NUS 3		64.250	12.276	50	12	30.00
NUS 4		67.250	12.657	60	15	37.50
NUM	B	70.425	10.129	60	10	22.50
NUS 1	C	66.875	11.076	60	15	37.50
NUS 2		65.125	9.439	60 & 70	12	30.00
NUS 3		65.625	10.512	60	12	30.00
NUS 4		64.750	11.263	60	15	37.50
NUM	C	69.850	9.181	70	12	30.00
NUS 1	D	58.250	16.777	60	20	50.00
NUS 2		60.000	17.394	60	15	37.50
NUS 3		65.500	11.536	70	18	45.00
NUS 4		70.000	10.127	60,70, 80	12	30.00
NUM	D	59.225	11.445	65	10	25.00

Dari hasil NUS dan NUM di empat kelas diketahui bahwa modus NUS adalah 60, kecuali kelas A. Sedangkan modus NUM adalah 70 untuk kelas A dan C. Untuk range modus NUS dan NUM antara 60-80.

Tabel 4. Korelasi NUS dan NUM Matematika di SMP N 4 Bandar Baru, Kabupaten Pidie

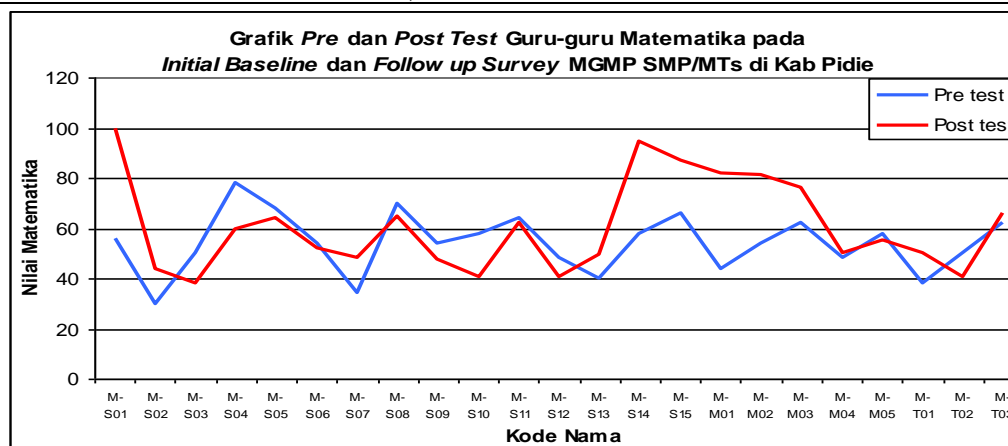
Kelas	Output	NUS Terbaik				NUM	Korelasi Karl - Pearson				r tabel
		1	2	3	4		1 dan NUM	2 dan NUM	3 dan NUM	4 dan NUM	
A	Mean	77.25	74.6	72.125	70.7	75.95	0.769	0.826	0.92	0.924	0.413
	Stdv	7.6753	7.8178	7.5556	7.5962	8.68819					
B	Mean	67.25	66.775	66.575	66.125	70.425	0.715	0.867	0.925	0.93	0.413
	Stdv	12.6567	10.6662	9.6527	9.8507	10.12964					
C	Mean	66.875	66.325	65.283	65.775	69.85	0.786	0.917	0.923	0.957	0.413
	Stdv	11.0759	9.23868	8.8633	8.7222	9.18067					
D	Mean	70	67.75	65.1	63.5	59.225	0.737	0.76	0.889	0.888	0.413
	Stdv	10.1274	9.19518	10.767	11.94	11.4455					

Tabel 5. Pre-Post Test dari Baseline dan Follow Up Survey MGMP Matematika SMP/MTs di Kabupaten Pidie.

No	Kabupaten/Lokasi	Kode	Total Pre test	Rata-rata Pre test	Total Post test	Rata-rata Post test
		Nama				
1	Pidie/Sigli	M-01	56	55.20	100.00	59.49

2		M-02	30		43.60	
3		M-03	50		38.40	
4		M-04	78		60.00	
5		M-05	68		64.40	
6		M-06	54		52.00	
7		M-07	34		48.00	
8		M-08	70		64.80	
9		M-09	54		47.60	
10		M-10	58		40.40	
11		M-11	64		62.00	
12		M-12	48		40.40	
13		M-13	40		49.60	
14		M-14	58		94.40	
15		M-15	66		86.80	
16	Pidie/Meureudu	M-01	44	53.20	82.00	68.88
17		M-02	54		81.20	
18		M-03	62		76.00	
19		M-04	48		50.00	
20		M-05	58		55.20	
21	Pidie/Tangse	M-01	38	50.00	50.00	52.13
22		M-02	50		40.40	
23		M-03	62		66.00	

(Sumber : LBB-BIS Banda Aceh, 2006/2007)



Dari hasil *pre* dan *post test* untuk guru-guru matematika SMP/MTs di Kabupaten Pidie di tiga lokasi yang berbeda, secara umum diperoleh modus total *score* 50, walaupun secara *real* ada yang memperoleh nilai 100. Sedangkan nilai rata-rata *pre test* adalah 54.09 dan nilai rata-rata *post test* adalah 60.57. Hal ini menunjukkan bahwa terdapat korelasi antara perolehan nilai rata-rata guru matematika dengan modus (*Mo*) NUS dan NUM siswa pada mata pelajaran matematika berbobot sekitar 60 (lihat Tabel 3, 4, 5 dan Grafik di atas). Kondisi ini menunjukkan adanya korelasi searah antara guru dan siswa, sehingga memberikan informasi bahwa pembinaan dan peningkatan konten (*capacity building*) untuk guru-guru SMP/MTs dapat menjadi prioritas utama, searah

dengan usaha meningkatkan hasil belajar siswa SMP/MTs untuk mata pelajaran matematika di Kabupaten Pidie. Artinya perolehan angka hasil evaluasi siswa (*Mo*) dan hasil evaluasi guru-guru matematika dalam MGMP memiliki bobot yang dapat dikategorikan sama. Oleh karena itu, sudah sepatutnya peningkatan mutu pembelajaran matematika dimulai dari guru-guru sebagai perantara transformasi dan transfer ilmu kepada siswa. Bila hasil transformasi dan transfer ini sama, maka dapat dikatakan bahwa perolehan bobot yang dihasilkan searah.

Tabel 6. *Check List* Pembelajaran Matematika di SLTP N 4 Bandar Baru-Pidie.

	Kelas 3 SLTPN 4 Bandar Baru, Kab. Pidie			
	A	B	C	D
<b>A. Faktor Utama</b>				
A1. Korelasi antara Ujian Sumatif dan Ulangan Umum	0,924	0,93	0,957	0,888
<b>B. Faktor Pendukung</b>				
<b>B1. Siswa</b>				
B1.1. a. Karya siswa ditulis menggunakan bahasa sendiri	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
b. Karya siswa ditulis dengan menyalin di papan tulis dan bahasa sendiri	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
B1.2. Karya siswa tidak dipajang	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<b>B2. Guru</b>				
B2.1. a. Guru membawa siswa belajar di luar kelas	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
b. Guru membawa siswa belajar di luar kelas dan bahan belajar dari lingkungan sekitar	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
B2.2. Guru berkeliling di kelas dan memberi komentar terhadap pekerjaan siswa	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
B2.3. a. Guru memiliki strategi untuk mendorong siswa mengajukan pertanyaan	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
b. Guru memiliki strategi untuk mendorong siswa untuk mengajukan pertanyaan dan mengemukakan pendapat	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
B2.4. Guru memberi kisi-kisi (bahan/bab), memberi tugas/PR, memberitahukan ujian sebelumnya	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
B2.5. a. Tes tertulis, bertanya untuk menguji pemahaman siswa (perorangan), mengamati tugas/pekerjaan siswa	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
b. Tes tertulis, tanya jawab langsung ke seluruh siswa, bertanya untuk menguji pemahaman siswa (perorangan), mengamati tugas/pekerjaan siswa	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Tabel 6. lanjutan

B2.6. a. Tes tertulis = 45 menit	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
b. Tanya jawab langsung ke seluruh siswa = 5 menit	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
c. Tanya jawab langsung ke seluruh siswa = 10 menit	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
d. Bertanya untuk menguji pemahaman siswa (perorangan) = 10 menit	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
e. Bertanya untuk menguji pemahaman siswa (perorangan) = 15 menit	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
B2.7. a. Tes tertulis esai dan kombinasi esai dan pilihan ganda	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

b. Tes tertulis esai , pilihan ganda dan kombinasi	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
B2.8. a. Sama model dan jumlah soal ujian sumatif, esai dengan jumlah 10	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
b. Tidak sama model dan jumlah soal ujian sumatif, tanya jawab dan esai dengan jumlah 5	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
B2.9. a. Ujian sumatif 4 kali	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
b. Ujian sumatif 5 kali	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
B2.10. a. Umpan balik diberikan secara lisan kepada seluruh siswa	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
b. Umpan balik diberikan secara lisan keseluruhan siswa dan perorangan	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
B2.11. Guru menulis komentar terhadap karya siswa	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
B2.12. Hasil ujian diumumkan saat di kelas	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
B2.13. a. Tugas berasal dari buku panduan	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
b. Tugas berasal dari buku panduan, tetapi dikembangkan guru	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
c. Tugas berasal dari buku panduan, buku panduan tetapi dikembangkan guru, tugas menuntut siswa untuk mengemukakan ide-ide	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<b>B3. Kelas</b>				
B3.1. Tidak tersedia bahan pembelajaran matematika di kelas	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
B3.2. Pengaturan kursi dalam baris	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<b>B4. Orang tua/masyarakat</b>				
B4.1. Ada dukungan dari orang tua/masyarakat selama belajar matematika di rumah	37.5 %	75%	62.5 %	25%

#### 4. Kesimpulan dan Saran

Dari hasil penelitian ini ada beberapa hal yang dapat disimpulkan, yaitu:

- (1). Evaluasi yang digunakan guru matematika dalam mengevaluasi belajar siswa SMPN 4 Bandar Baru-Pidie tahun pelajaran 2006/2007 sudah valid secara kuantitatif, dengan nilai korelasi untuk kelas A sebesar 0.924, kelas B sebesar 0.93, kelas C sebesar 0.957, kelas D sebesar 0.888, dimana ini hanya melihat nilai-nilai, tetapi bukan melihat bobot soal-soal.
- (2). Dari keempat korelasi yaitu korelasi antara satu, dua, tiga dan empat ujian sumatif yang terbaik dan ulangan umum, hanya korelasi antara empat ujian sumatif yang terbaik dan ulangan umum yang korelasinya paling baik dari korelasi yang lain. Kecuali kelas D yang korelasi antara tiga ujian sumatif yang terbaik dan ulangan umum yang nilai korelasinya terbaik dari korelasi yang lain, karena kelas D ada siswa yang tidak mengikuti ujian sumatif.
- (3). Tingkat validitas evaluasi yang digunakan guru matematika dalam menilai hasil belajar siswa di SMPN 4 Bandar Baru-Pidie tahun 2006/2007 memiliki tingkat validitas tinggi secara kuantitatif bagi kelas III D dan kelas III A, B, C memiliki tingkat validitas sangat tinggi (lihat [9]).
- (4). Adanya korelasi searah antara evaluasi belajar siswa dan evaluasi *initial baseline* dan *follow up survey* MGMP matematika tingkat SMP/MTs di Kabupaten Pidie.

Berdasarkan hasil penelitian di atas maka ada beberapa hal yang dapat disarankan :

- (1). Perlu diperhatikan validitas secara kualitatif melalui materi yang diujikan, seperti tingkat kesulitan soal, penalaran, soal analisis, dan sebagainya dalam membuat alat evaluasi yang baik.
- (2). Diharapkan kepada guru-guru matematika, dalam membuat evaluasi matematika harus memperhatikan tingkat validitas tes.
- (3). Setiap guru-guru matematika agar setiap selesai mengadakan evaluasi, sebaiknya lembar soal tidak dibuang begitu saja, tetapi soal-soal tersebut dapat diteliti kembali untuk diambil item-item yang baik agar setiap guru memiliki bank soal.
- (4). Guru dapat melakukan tes tingkat validitas sendiri dengan tes korelasi sehingga diperoleh hasil yang optimal dari pelaksanaan suatu evaluasi.
- (5). Diharapkan kepada guru agar mempunyai strategi, metode, pendekatan, teknik secara kuantitatif dan kualitatif dalam mengevaluasi proses pembelajaran matematika.

### **Daftar Pustaka**

- [1] Akmal, 2003, <http://www.damandiri.or.id/file/akmalbab3.pdf>. 31 Juli 2007.
- [2] S. Arikunto, 1993. *Prosedur Penelitian Suatu Pendekatan Praktik*. PT Bina Aksara, Jakarta.
- [3] O. Hamalik, 2002. *Perencanaan Pengajaran Berdasarkan Pendekatan Sistem*. Bumi Aksara, Jakarta.
- [4] M. I. Hasan, 1999. *Pokok-pokok Materi Statistik 1*. Bumi Aksara, Jakarta.
- [5] H. Budi, 2000. *Aplikasi SPSS: Mengenal Tabel Frekuensi dan Tabulasi Silang*, [http://library.gunadarma.ac.id/files/disk1/1/jbptgunadarma-gdl-course-2000-budiherman-8-spss\\_fre-b.doc](http://library.gunadarma.ac.id/files/disk1/1/jbptgunadarma-gdl-course-2000-budiherman-8-spss_fre-b.doc). 18 Juni 2007.
- [6] R. Novita, 2007. *Validitas Dari Evaluasi Belajar Yang Digunakan Guru Matematika Dalam Menilai Hasil Belajar Siswa Di SLTP Negeri 4 Bandar Baru-Pidie*. Skripsi. Jurusan Matematika FMIPA Unsyiah.
- [7] W. Nurkencana, 1981, *Evaluasi Pendidikan*, Bina Aksara, Jakarta.
- [8] N. Purwanto, 1982, *Prinsip-prinsip dan Teknik Evaluasi Pengajaran*, Remaja Karya, Bandung.
- [9] Sudijono, A., 2001, *Pengantar Statistik Pendidikan*, PT. Rajagrafindo, Jakarta.
- [10] A. Suyuti, 2002. <http://www.damandiri.or.id/file/ahmadsuyutiunairbab4.pdf> [3Maret 2007]
- [11] M.E. Yusnandar, dan S. Sri, 2006, *Aplikasi Analisis Khi Kuadrat ( $X^2$ ) Terhadap Kekurangan Energi Protein pada Anak Dibawah Lima Tahun (Balita) dan Faktor-Faktor yang Berhubungan*, <http://www.litbang.deptan.go.id/warta-ip/pdf-file/5.yusipvol-15.pdf>. [31 Juli 2007]