

UJI VALIDITAS KONSTRUK TES PRESTASI BELAJAR DENGAN PENERAPAN MODEL PERSAMAAN STRUKTURAL

Oleh

Purwo Susongko

Universitas Pancasakti Tegal

Kusumatirto@gmail.com

Validitas konstruk adalah sejauh mana hasil pengukuran dianggap mencerminkan konstruk (*construct*) dalam suatu teori. Uji validitas konstruk pada tes prestasi belajar diperlukan untuk konfirmasi antara validitas konten/isi yang telah dibuat dan data empirik. Model analisis yang cukup komprehensif untuk uji validitas konstruk adalah analisis faktor yang terdiri dari analisis faktor eksplanatori dan analisis faktor konfirmatori. Analisis faktor konfirmatori dapat dilakukan dengan model persamaan struktural (*structural equation modeling, SEM*). Dengan Analisis konfirmatori peneliti dapat memodifikasi model untuk disesuaikan dengan data yang dimilikinya, sesuatu yang tidak dapat dilakukan pada analisis eksplanatori. Dengan menggunakan data empirik, terbukti model persamaan struktural sangat tepat digunakan untuk menguji validitas konstruk tes prestasi belajar di sekolah.

Kata kunci: Validitas konstruk, Analisis faktor konfirmatori, model persamaan struktural, tes prestasi belajar

A. Pendahuluan

Kegiatan pengukuran dan penilaian adalah merupakan upaya pengumpulan informasi yang benar dari para siswa. Informasi tersebut selanjutnya akan digunakan sebagai dasar dalam pengambilan keputusan-keputusan kebijaksanaan baik secara lokal maupun nasional. Agar dapat diperoleh informasi yang benar dan akurat yang sangat besar pengaruhnya dalam pengambilan keputusan, maka diperlukan alat-alat pengukur yang baik yaitu yang memenuhi syarat-syarat baik kesahihannya/validitas maupun keandalannya/reliabilitas. Disamping kedua syarat utama tersebut suatu alat ukur juga harus memenuhi kriteria yang lain seperti : obyektivitas, praktibilitas, reproduksibilitas, diskriminasi, dan komprehensif.

Kesalahan dalam pengukuran pendidikan menurut sifatnya dapat dibedakan dalam tiga jenis yaitu kesalahan yang sistematis, kesalahan yang bersifat bias dan kesalahan acak/random. Sedangkan kesalahan pengukuran berdasarkan sumbernya dapat berasal dari instrumen, peserta tes, penguji, maupun pelaksanaannya. Kesalahan sistematis adalah kesalahan pengukuran yang berlaku bagi semua peserta tes, sedangkan kesalahan acak adalah kesalahan yang tidak diketahui sebabnya dan berlaku bagi semua peserta tes, sedangkan

kesalahan bias adalah kesalahan pengukuran yang menguntungkan atau merugikan bagi sekelompok peserta tes. Validitas adalah ukuran kecermatan suatu tes dari kesalahan sistematis, sedangkan reliabilitas adalah kecermatan suatu tes dari kesalahan acak. Tes yang baik harus tersusun dari item-item yang baik, oleh karenanya disamping ada ukuran validitas dan reliabilitas tes, dikenal pula validitas item. Ukuran validitas item pada umumnya menggunakan tingkat kesukaran item, daya beda item dan tingkat tebakan menjawab benar (*guessing*).

Validitas didefinisikan sebagai ukuran seberapa cermat suatu alat ukur melakukan fungsi ukurnya. Alat ukur yang valid (sahih) adalah alat ukur yang mampu mengukur apa yang akan diukur atau yang dapat memenuhi fungsinya sebagai alat ukur. Jadi, untuk dikatakan valid, tes harus mengukur sesuatu dan melakukannya dengan cermat (Mardapi, 2004). Validitas suatu alat ukur sangat erat kaitannya dengan kelompok yang akan diukur, hal ini disebabkan karena memang hanya berlaku pada kelompok yang bersangkutan atau kelompok lain yang kondisinya sama. (Sumadi Suryabrata, 1983).

Lebih lanjut, pengertian validitas suatu tes tidaklah berlaku umum untuk semua tujuan ukur. Sebuah tes biasanya hanya menghasilkan ukuran yang valid untuk satu tujuan ukur tertentu. Karena itu predikat valid seperti dalam pernyataan "Tes ini valid" tidaklah benar. Pernyataan valid harus diiringi oleh keterangan yang menunjuk kepada tujuan, yaitu valid untuk mengukur apa dan untuk siapa. Validitas juga digambarkan sebagai suatu penetapan evaluasi terintegrasi tentang derajat bukti empiris dan dasar teoretis yang mendukung ketercukupan dan kesesuaian tindakan dan kesimpulan yang berdasarkan pada skor tes atau model-model lain dari penilaian (Messick, 2001). Atkin, Black & Cofey (2001) berpendapat validitas dapat ditinjau dari dimensi validitas isi (*content validity*), validitas konstruk (*construct validity*), dan validitas instruksional (*instructional validity*) (Mansyur, Rasyid & Suratno, 2009). Sedangkan *American Psychological Association* membagi validitas menjadi *content validity*, *construct validity*, dan *criterion-related validity*.

1. *Content – Validity*

Suatu alat ukur dianggap memiliki validitas isi, apabila alat ukur tersebut isinya telah dapat mengukur dari keseluruhan isi dari yang akan diukur. Dengan demikian keputusan yang didasarkan atas validitas isi menentukan apakah siswa telah menguasai atau mahir, atau gagal

dalam menjawab butir-butir soal yang mengukur tujuan belajar sebagaimana yang termuat dalam kurikulum. Validitas isi digunakan bila hal yang akan diukur diketahui dengan jelas dan komprehensif seperti halnya pada tes prestasi belajar.

Dengan perkataan lain suatu tes memiliki validitas isi bilamana tingkah laku dan isi pelajaran yang dikandung dalam butir soal berhubungan dengan tingkah laku dan isi pelajaran yang dinyatakan dalam tujuan khusus, kandungan butir soal yang tidak berhubungan dengan tujuan khusus (sasaran belajar akan menyebabkan isi tes tidak valid). Terdapat dua macam tipe validitas isi yaitu *face validity* dan *logical validity*. *Face validity* tercapai apabila pemeriksaan terhadap item-item tes memberi kesimpulan bahwa tes tersebut mengukur aspek yang relevan. Dasar penyimpulannya lebih banyak diletakan pada akal sehat/*cammon sense*. *Face validity* adalah tipe validitas yang paling rendah signifikansinya. *Logical validity* disebut juga *sampling validity*. Validitas ini menuntut batasan yang seksama terhadap kawasan(domain) perilaku yang diukur dan suatu desain logis yang dapat mencakup bagian-bagian kawasan perilaku tersebut. Sejauhmana tipe validitas ini telah terpenuhi dapat dilihat dari cakupan butir-butir yang ada dalam tes.validitas isi dapat dilakukan dengan : (1) membuat kisi-kisi soal atau spesifikasi tes, (2) meminta pendapat pakar/ahli berkaitan dengan isi tes.

2. *Criterion-related validity*

Criterion-related validity adalah validasi yang didasarkan adanya korelasi antara alat ukur yang sedang kita buat dengan alat ukur lain yang sudah diketahui valid (sebagai kriterium). Jenis validitas ini dapat dibedakan menjadi dua yaitu :

a. *Predictive-validity* (validitas peramalan) yaitu apabila pengukuran yang dilakukan dipergunakan untuk memprediksi atau meramal tingkah laku atau prestasi yang akan datang. Contoh suatu alat ukur yang mendasarkan pada validitas peramalan (*predictive-validity*) adalah tes masuk perguruan tinggi, sekolah maupun tes seleksi pegawai. Mereka yang mendapat skor baik pada tes masuk Perguruan tinggi diharapkan akan memiliki prestasi yang tinggi pula setelah kuliah. Bila kenyataan tersebut tercapai maka tes masuk tersebut memiliki validitas peramalan yang meyakinkan. Sebagai contoh bila ingin mengetahui validitas prediktif tes ujian masuk sekolah maka secara empirik dilakukan dengan mengkorelasikan skor ujian masuk tersebut dengan skor ujian setelah siswa mengikuti

program pembelajaran di sekolah. Bila korelasi signifikan artinya tes seleksi masuk sekolah tersebut valid sebagai tes seleksi.

b. *Concurrent-validity* (validitas sama saat): adalah validitas yang didasarkan pada kriterium yang diperoleh pada waktu yang bersamaan atau hampir bersamaan dengan prediktor (tes yang sedang dibuat). Misalnya seorang guru ingin menggunakan tes tentang keterampilan belajar untuk mengestimasi hasil pengamatan yang diperoleh dalam situasi belajar yang sebenarnya. Kedua pengukuran tersebut baik mengenal “keterampilan belajar” maupun “kegiatan belajar” yang sebenarnya dilaksanakan secara bersamaan. Kedua hasil pengukuran tersebut kemudian dikorelasikan, dan hasilnya adalah merupakan koefisien validitas tes keterampilan belajar. Prosedur pengujian validitas ini dinamakan “*Concurrent Validity*”.

3. *Construct – Validity*

Adalah sejauh mana hasil pengukuran dianggap mencerminkan konstruk (*construck*) dalam suatu teori. Konstruk pada dasarnya adalah bagian variabel yang dapat diukur, oleh karenanya konstruk diturunkan dari definisi konseptual dari suatu variabel atau dapat dikatakan konstruk adalah definisi operasional dari variabel. Oleh karenanya, maka validitas konstruk dibutuhkan untuk tes-tes dimana bahan yang akan diujikan masih perlu didefinisikan terlebih dahulu. Pengujian validitas konstruk pada dasarnya menguji secara empirik terhadap definisi operasional yang digunakan sebagai dasar penyusunan instrumen. Contoh : “*construct validity*” dapat dijelaskan sebagai berikut : Seorang guru telah menemukan suatu teori tentang “motivasi belajar”. Selanjutnya dia berkeinginan menyusun alat ukur yang dapat mengukur motivasi siswa. Dengan mendasarkan pada teori motivasi yang dia ketahui maka guru tersebut berhasil menyusun alat ukur motivasi. Alat ukur motivasi yang dibuat oleh guru tadi dianggap valid kalau dapat mengungkap perbedaan-perbedaan antara individu satu dengan yang lain dilihat dari motivasi mereka.

Walaupun tes prestasi belajar pada dasarnya telah mempunyai domain yang jelas, namun tidak berarti tidak dibutuhkan lagi uji terhadap validitas konstruksya. Uji validitas konstruk pada tes prestasi belajar diperlukan untuk konfirmasi antara validitas konten/isi yang telah dibuat dan data empirik setelah dilakukan uji coba. Untuk tes prestasi belajar validitas konten biasanya dituangkan dalam kisi-kisi soal. Kisi-kisi soal tersebut disusun berdasarkan

standar kompetensi, kompetensi dasar dan indikator yang telah dibuat oleh guru atau asesor independen. Standar kompetensi yang disusun berdasarkan standar kelulusan yang telah dirumuskan baik standar nasional bahkan mungkin standar internasional.

Proses pengujian “*Construct validity*” dapat dilakukan sebagai berikut :

- a. Mengidentifikasi secara rinci tentang . “*construct*” yang akan memberikan jawaban hasil tes.
- b. Menyusun hipotesis-hipotesis yang berhubungan dengan “*construct*”.
- c. Mengumpulkan data untuk menguji hipotesis tersebut.
- d. Menganalisa data dan mengumpulkan.

Model analisis yang cukup komprehensif untuk uji validitas konstruk adalah analisis faktor yang terdiri dari analisis faktor eksplanatori dan analisis faktor konfirmatori. Analisis konfirmatori dapat dilakukan dengan model persamaan struktural (*structural equation modeling, SEM*). Secara sederhana validitas konstruk dapat dilakukan dengan mengkorelasikan skor tiap item dengan skor total untuk suatu perangkat tes yang telah diujicobakan terlebih dulu.

Tulisan ini akan mengkaji beberapa hal sebagai berikut: (1) Membahas konsep dasar Validitas dan hubungannya dengan kesalahan pengukuran dalam pengukuran pendidikan , (2) Membahas analisis faktor konfirmatori sebagai bagian dari model persamaan struktural sebagai salah satu metode dalam menguji validitas konstruk tes prestasi belajar, (3) Menguji validitas konstruk dengan model persamaan struktural menggunakan data empirik tes prestasi belajar Kimia di SMA.

B. Model Persamaan Struktural

Model persamaan struktural (SEM) memiliki dua bagian penting yaitu model pengukuran (measurement model) yang dilaksanakan dengan analisis konfirmatori dan model struktural (structural model) yang dilakukan dengan model regresi. Bagian model pengukuran ini berfungsi mengetahui sejauh mana ketepatan pengukuran yang telah dilakukan oleh peneliti sedangkan model structural digunakan untuk menguji suatu teori atau hubungan antar konsep. Teknik SEM menjadi unggul karena dalam SEM dapat dilakukan kedua langkah tersebut secara bersama sehingga tidak diperlukan lagi uji coba instrumen pada

populasi lain yang sejenis dengan populasi penelitian. Kedua, pada analisis konfirmatori peneliti dapat memodifikasi model untuk disesuaikan dengan data yang dimilikinya, sesuatu yang tidak dapat dilakukan pada analisis eksplanatori. Ketiga, dengan analisis konfirmatori dapat pula ditentukan reliabilitas tiap konsep maupun indikator secara bersama dan komprehensif.

SEM terdiri atas beberapa tahapan sebagai berikut (Widodo, 2006) :

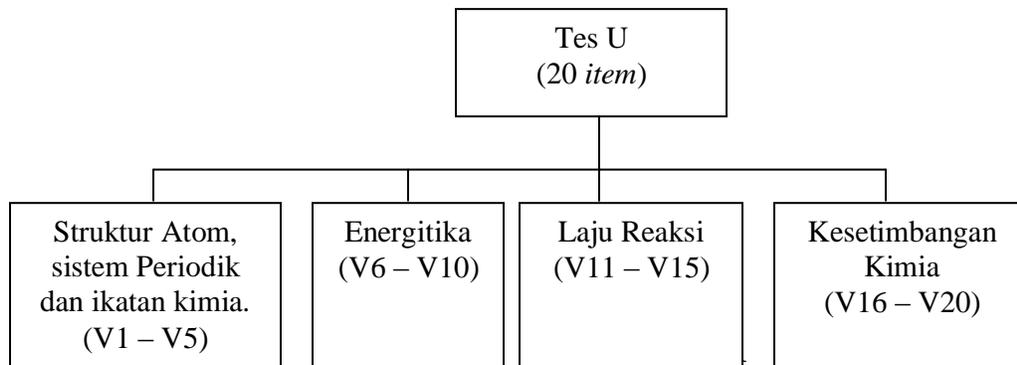
1. *Pengembangan model berdasarkan teori*: Tujuannya adalah untuk mengembangkan sebuah model yang mempunyai justifikasi (pembenaran) secara teoritis yang kua guna mendukung upaya analisis terhadap suatu masalah yang sedang dikaji/diteliti.
2. *Pengembangan diagram lintasan (path diagram)*:Tujuannya adalah menggambarkan model teoritis yang telah dibangun pada langkah pertama kedalam sebuah diagram jalur agar peneliti dengan mudah dapat mencermati hubungan kausalitas yang ingin diujinya.
3. *Mengkonversi diagram jalur kedalam persamaan structural*:Langkah ini membentuk persamaan-persamaan pada model struktural dan model pengukuran
4. *Pemilihan data input dan teknik estimasi*:Tujuannya adalah menetapkan data input yang digunakan dalam pemodelan dan teknik estimasi model
5. *Evaluasi masalah identifikasi model*:Tujuannya adalah untuk mendeteksi ada tidaknya masalah identifikasi berdasarkan evaluasi terhadap hasil estimasi yang dilakukan program komputer
6. *Evaluasi Asumsi dan Kesesuaian model*:Tujuannya adalah untuk mengevaluasi pemenuhan asumsi yang disyaratkan SEM, dan kesesuaian model berdasarkan kriteria *goodness-of-fit tertentu*.
7. *Interpretasi dan modifikasi model*:Tujuannya adalah untuk memutuskan bentuk perlakuan lanjutan setelah dilakukan evaluasi asumsi dan uji kesesuaian model.

Penerapan model SEM pada tes prestasi belajar terutama pada tes-tes yang mengandung beberapa standar kompetensi seperti halnya ujian akhir semester, ujian

kenaikan kelas, ujian sekolah maupun ujian nasional, sangat lah diperlukan. Hal ini disebabkan dengan penerapan model SEM, dapat diketahui: (1) Signifikansi dari indikator, kompetensi dasar , atau standar kompetensi terhadap kompetensi pokok yang diharapkan, (2) Eksistensi dan urutan suatu standar kompetensi atau kompetensi dasar dalam sumbangnya terhadap kompetensi pokok yang diharapkan, (3) Evaluasi terhadap kurikulum.

C.Kajian Empirik Pada Tes Prestasi Belajar Kimia

Tes Prestasi belajar kimia dalam bentuk uraian di konstrak atas empat pokok bahasan yaitu: (1) struktur atom, sistem periodik, dan ikatan kimia; (2) energitika; (3) laju reaksi; dan (4) kesetimbangan kimia. Pokok bahasan tersebut diberikan pada siswa kelas XI SMA semester gasal. Dari keempat pokok bahasa tersebut diuraikan dalam 20 *item* dengan masing-masing *item* mempunyai 4 kategori yaitu *item* yang mempunyai langkah penyelesaian sebanyak 3 langkah.. Rancangan tes dapat digambarkan sebagai berikut :



Gambar 1

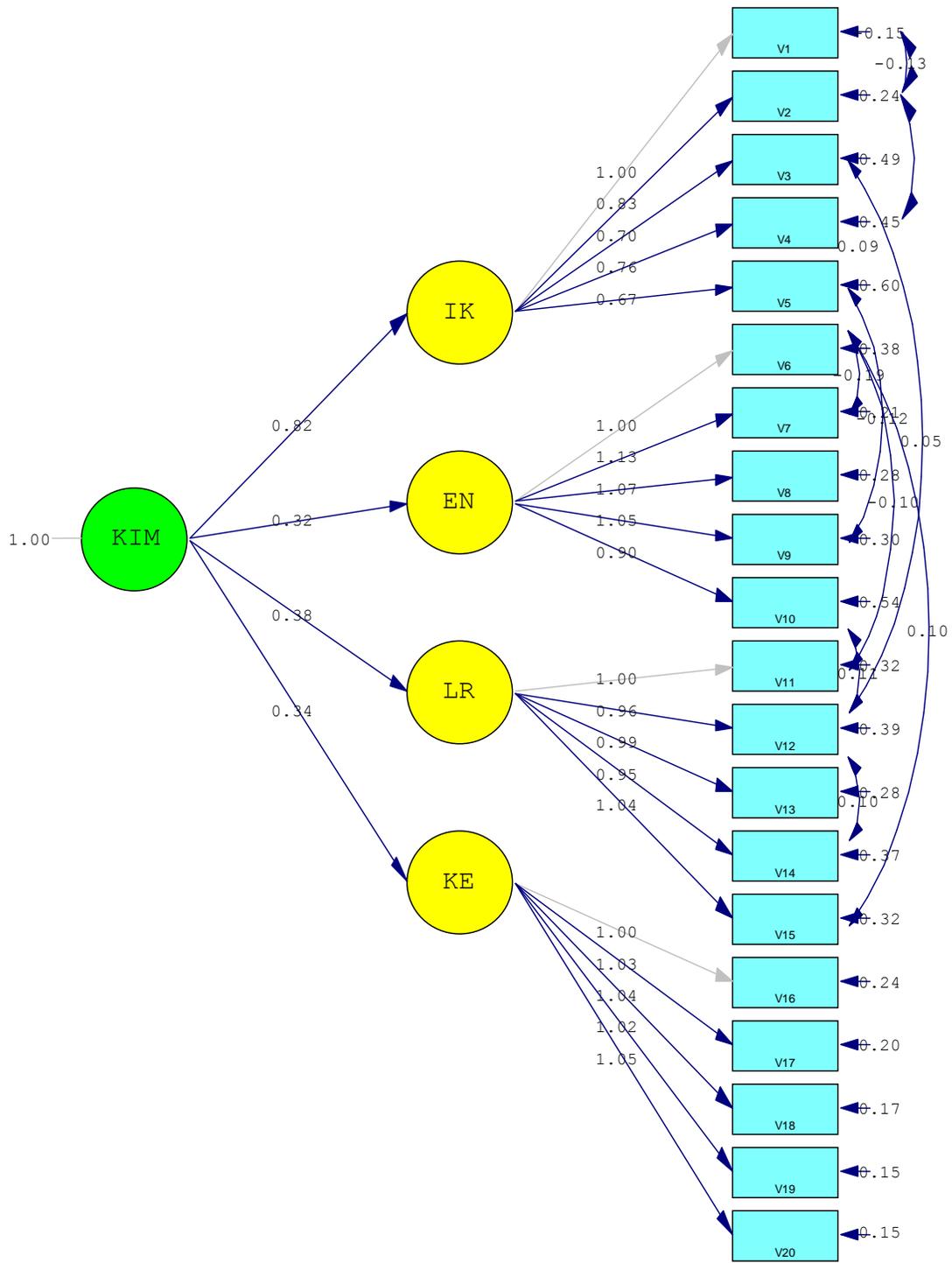
Konstrak Tes U

Sebelum dilakukan uji coba, tes di telaah oleh satu orang ahli pendidikan kimia dan dua orang guru senior yang telah lebih dari 20 tahun mengajar kimia di SMA. Hal ini dilakukan untuk memenuhi validitas rasional sebelum dibuktikan secara empirik, Uji coba dilaksanakan untuk memperoleh bukti validitas konstrak. Tes diujicobakan kepada 253 siswa kelas XI Program IPA SMA yang diambil dari 2 SMA yaitu SMA 1 SLAWI dan SMA 3 SLAWI, Kabupaten Tegal. Penskoran dilakukan dengan model analitik secara terurut sehingga pengerjaan soal harus melalui tahapan secara terurut. Ada tiga *rater* dalam penskoran tes ini yaitu dua guru senior dan peneliti. Bila dari dua guru senior terjadi perbedaan skor, barulah peneliti menentukan skor setelah terlebih dahulu berdiskusi dengan dua guru senior maupun peneliti.

Untuk membuktikan validitas konstruk dan asumsi keunidimensian suatu tes, data yang diperoleh dari hasil uji coba dianalisis menggunakan Model Persamaan Struktural (MPS). MPS yang digunakan adalah Analisis Faktor Konfirmatori (AFK).

Ada beberapa ukuran *goodness of fit* yang digunakan dalam menilai ketepatan dari model yang telah disusun. Hal ini dilakukan agar didapatkan dukungan empirik yang komprehensif terhadap model. Ukuran-ukuran tersebut adalah RMSEA, NFI, NNFI, CFI, IFI, RFI dan PGFI.

Model analisis faktor konfirmatori yang diperoleh dari hasil analisis di tunjukkan dalam Gambar 2. Dukungan terhadap model yang dikembangkan dari data empirik (sampel), dapat dilihat dari besarnya nilai RMSEA (*root mean square error of approximation*). RMSEA ini mengukur penyimpangan nilai parameter pada suatu model dengan matriks kovarian populasi (Imam Ghazali & Fuad, 2005: 31). Mac Callum et al (1996) menyatakan bahwa RMSEA berkisar antara 0,08 sampai dengan 0,1 merupakan model yang memiliki fit cukup, yaitu memiliki nilai RMSEA sebesar 0,083. Dukungan terhadap model yang lain dapat dilihat dari nilai NFI (*Normed Fit Index*) sebesar 0,90 (*cut-off* sebesar 0,9), NNFI (*Non – Normed Fit Index*) sebesar 0,92 (*cut-off* sebesar 0,9), CFI (*Comparative Fit Index*) sebesar 0,93 (*cut-off* sebesar 0,9) dan IFI (*Incremental Fit Index*) sebesar 0,93 (*cut-off* sebesar 0,9), RFI (*Relative Fit Index*) sebesar 0,90 (mendekati 1), PGFI (*Parsimony Goodnes Of Fit Index*) sebesar 0,64 (*cut-off* sebesar 0,6) (Imam Ghazali & Fuad, 2005: 29-33).



Chi-Square=428.50, df=157, P-value=0.00000, RMSEA=0.083

Gambar 2
Model Uji Validitas Konstrak Tes Prestasi Kimia

Berdasarkan hasil tersebut, validitas konstruk tes terbukti kebenarannya. Namun demikian, estimasi terhadap parameter perlu diuji satu per satu, apakah hasilnya signifikan secara statistik. Koefisien estimasi hubungan antara variabel laten struktur atom, sistem periodik dan ikatan kimia (IK), energetika (EN), laju reaksi (LR) dan kesetimbangan kimia (KE) dengan variabel-variabel indikatornya secara statistik signifikan. Besarnya nilai koefisien estimasi (λ), nilai t dan *standard error* yang dihasilkan pada analisis di atas, ditunjukkan oleh Tabel 1.

Tabel 1 menggambarkan pengaruh variabel laten struktur atom, sistem periodik dan ikatan kimia (IK) terhadap indikator x1 – x5 secara statistik signifikan, energetika (EN) terhadap indikator x6 – x10 secara statistik signifikan, laju reaksi (LR) terhadap indikator x11 – x15 secara statistik signifikan dan kesetimbangan kimia (KE) terhadap indikator x16 – x20 secara statistik signifikan .

Tabel 1 menunjukkan bukti bahwa variabel laten struktur atom, sistem periodik dan ikatan kimia (IK) secara teoritis dapat diukur dengan indikator x1 – x5, energetika (EN) secara teoritis dapat diukur dengan indikator x6 – x10, laju reaksi (LR) dapat diukur dengan indikator x11 – x15 dan kesetimbangan kimia (KE) secara teoritis dapat diukur dengan indikator x16 – x20. Selanjutnya, hubungan antara variabel laten (K, EN, LR,KE) dengan variabel KIM pada bentuk tes uraian dapat digambar pada Tabel 2

Tabel 1

Koefisien λ , *Standard Error*, dan Nilai t pada Tes U

Var	struktur atom, sistem periodik dan ikatan kimia (IK)	Va	energetika (EN)	Var	laju reaksi (LR)	Va	kesetimbangan kimia (KE)
X1	1,000	X6	1,000	X11	1,000	X16	1,000
X2	0,83 (0,038) 21,91*	X7	1,13 (0,062) 18,20*	X12	0,96 (0,063) 15,19*	X17	1,03 (0,045) 22,67,13*
X3	0,70 (0,037) 19,15*	X8	1,07 (0,062) 16,36*	X13	0,99 (0,058) 17,18*	X18	1,04 (0,046) 22,79*
X4	0,76 (0,034) 22,08*	X9	1,05 (0,066) 16,03*	X14	0,95 (0,062) 15,42*	X19	1,02 (0,043) 24,05*
X5	0,67 (0,034) 19,62*	X10	0,90 (0,069) 12,95*	X15	1,04 (0,061) 17,14*	X20	1,05 (0,045) 23,61*

Tabel 2

Nilai Koefisien Gamma (J), *Standard Error* dan Nilai t
pada Tes Bentuk Uraian

Variabel	Kim		
	Koef Gamma (λ)	<i>Standard Error</i>	Nilai t
IK	0, 82	0, 11	7, 31*
EN	0, 32	0, 069	4, 57*
LR	0, 38	0, 075	5, 08*
KE	0, 34	0, 075	4, 50*

* : Signifikan pada $\alpha = 0, 05$

D. Kesimpulan

Dari hasil analisis konfirmatori terbukti bahwa standar kompetensi untuk materi struktur atom, sistem periodik dan ikatan kimia (IK), energetika (EN), laju reaksi (LR) dan kesetimbangan kimia (KE) secara statistik signifikan sebagai prediktor untuk tes prestasi belajar kimia SMA program IPA kelas XI untuk semester gasal. Ini berarti Tes kimia) secara teoritis dapat diukur dengan variable laten, IK, EN, LR dan KE. Hasil analisis ini menunjukkan bahwa validitas konstruk terbukti .

F. Daftar Pustaka

Holland. PW & Thaycr. DT (1988). *Test Validity*. New Jersey: Lawrence Erlbaum Associates, Publishers

Imam Ghozali & Fuad, (2005). *Structural equation modeling*, Semarang: Badan Penerbit Universitas Diponegoro

Joreskog, K.G. & Sorbom, D. (1996). *LISREL 8: user's reference guide*. Chicago, IL:SSI, Inc.

Joreskog, K.G. & Sorbom, D., Toit,S. & Toit, M.. (2000). *LISREL 8: New statistical features.* Chicago, IL:SSI, Inc.

Kerlinger, Fred N (1996). *Asas-asas Penelitian Behavioral* (Edisi Ketiga), diterjemahkan Simatupang L. R. Yogyakarta: Gadjah Mada University Press.

Kusnendi.(2008). *Model-model Persamaan Struktural*. Bandung: CV Alfabeta

- Mansyur, Rasyid, H & Suratno.(2009). *Asesmen Pembelajaran di Sekolah*. Yogyakarta: Multi Pressindo
- Mardapi, D. (Maret 1997). *Ragam bentuk evaluasi*. Makalah Semiloka Evaluasi Sistem Penilaian dan Pengukuran Hasil Belajar Mahasiswa UGM, di Universitas Gadjah Mada .
- _____.(1999). *Estimasi kesalahan pengukuran dalam bidang pendidikan dan implikasinya pada ujian nasional*. Pidato Pengukuhan Guru Besar. Yogyakarta. 11 September 1999
- _____.(2004). *Penyusunan tes hasil belajar*. Tidak diterbitkan. Program Pascasarjana, Universitas Negeri Yogyakarta.
- _____.(April 2005). *Penerapan matematika dan statistika pada pengukuran pendidikan*. Makalah Seminar Nasional Matematika II, di Universitas Pancasakti .
- Retnowati, P.(2004). *Seribu Pena Kimia SMA untuk kelas XI*. Jakarta: Erlangga
- Setyo H Wijayanto. (2002). *Structural equation modeling & lisrel 8.51 for window*. Tidak diterbitkan. Jurusan akuntansi, Fakultas Ekonomi UI
- Suryabrata, S. (1984). *Pembimbing ke psikodiagnostik (edisi ke 2)*. Yogyakarta: Sarasin
- _____. (1987). *Pengembangan tes hasil belajar*. Jakarta: CV. Rajawali.
- Messick, Samuel. (1993). “*Validity*”, *Educational Measurement, Third Edition*, ed. Robert L. Linn. New York: American Council on Education and Macmillan Publishing Company, A Division of Macmillan, Inc.
- Mueller, RO.(1996). *Basic principle of structural equation modeling: An introduction to LISREL and EQS*. Springer-Verlag New York, Inc.