

PENGGUNAAN MODEL RASCH UNTUK ANALISIS INSTRUMEN TES PADA MATA KULIAH MATEMATIKA AKTUARIA

Azizah^{1*}, Sapti Wahyuningsih²

^{1,2}Jurusan Matematika, Universitas Negeri Malang
Jl. Semarang No.5, Sumbersari, Kec. Lowokwaru, Kota Malang, Jawa Timur, Indonesia

e-mail: ¹azizah.fmipa@um.ac.id; ²sapti.wahyuningsih.fmipa@um.ac.id;

*corresponding author**

Abstrak

Penelitian ini dilakukan untuk menganalisa instrumen tes yang digunakan untuk mengukur kemampuan mahasiswa pada mata kuliah matematika aktuarial di jurusan matematika Universitas Negeri Malang. Peserta mata kuliah ini terdiri dari 40 mahasiswa. Soal yang diberikan berupa soal pilihan ganda berjumlah 50 butir terkait materi perhitungan premi pada asuransi jiwa. Model Rasch digunakan untuk mendapatkan butir soal yang *fit*. Analisa ini dilakukan dengan bantuan *software Winsteps*. Dari output program *Winsteps* diperoleh hasil 25 butir soal sesuai dengan model Rasch dengan nilai rata – rata Outfit MNSQ untuk person dan item masing – masing 0.98 dan 0.98. Sedangkan nilai Outfit ZSTD untuk person dan item masing – masing 0 dan -0.01. Sedangkan reliabilitas instrumen yang dinyatakan dalam *alpha cronbach* bernilai 0.85.

Kata Kunci: *fit*, model rasch, *winsteps*, reliabilitas.

THE USE OF RASCH MODEL FOR ANALYZING TEST INSTRUMENT IN ACTUARIAL MATHEMATICS COURSE

Abstract

This research was conducted to analyze the test instruments used to measure the ability of students in actuarial mathematics courses in mathematics department, State University of Malang. Participants in this course consist of 40 students. The form of questions given were multiple choice questions totaling 50 items related to the premium calculation material on life insurance. The Rasch model is used to get fit items. This analysis is carried out with the help of *Winsteps* software. From the *Winsteps* program output, 25 items were obtained fit to the Rasch model with an average MNSQ Outfit score for person and item 0.98 and 0.98, respectively. While the Outfit ZSTD value for person and item are 0 and -0.01 respectively. While the reliability of the instrument stated in *alpha cronbach* is 0.85.

Keywords: *fit*, rasch model, *winsteps*, reliability.

1. Pendahuluan

Dalam pembelajaran di lingkungan kampus atau pendidikan tinggi, dosen memegang peranan penting dalam melakukan proses evaluasi serta penilaian kemampuan mahasiswa pada mata kuliah yang diajarkan. Proses evaluasi perlu dilakukan dengan baik agar mampu mengukur kemampuan mahasiswa yang sebenarnya. Penilaian yang baik membutuhkan instrumen tes yang baik pula. Penilaian hasil pembelajaran yang dilakukan oleh dosen dapat menjadi langkah penting guna memperbaiki proses pembelajaran. Selain itu, hal ini dapat berfungsi untuk meningkatkan kualitas

belajar para mahasiswa atau peserta didik. Agar tujuan – tujuan tersebut dapat tercapai, proses evaluasi perlu dilaksanakan secara berencana, bertahap dan terus menerus untuk memperoleh gambaran perkembangan belajar dari para mahasiswa.

Menurut (Dunn, Morgan, O'Reilly, & Parry, 2003. p. 16), terdapat tiga tujuan proses penilaian dalam proses pembelajaran yaitu mendiagnosis kesulitan belajar peserta didik, mengukur peningkatan dari waktu ke waktu dan menggali informasi yang dapat digunakan oleh peserta didik untuk meningkatkan prestasi mereka. Selama ini,



proses penilaian pembelajaran yang dilakukan oleh dosen kepada peserta didik sangat berkaitan dengan proses perolehan atau pengumpulan data kemampuan belajar mahasiswa sehingga dapat diketahui kemampuan seorang mahasiswa dibandingkan mahasiswa yang lain. Selain itu penilaian pembelajaran juga dapat digunakan untuk mengetahui keberhasilan dosen dalam proses pembelajaran di kelas. Pada prakteknya, pengukuran kompetensi tertentu selama proses pembelajaran dapat dilakukan dengan metode tes. Pemilihan metode tes yang tepat diharapkan mampu mengukur kemampuan peserta didik dengan baik.

Pada penelitian ini pengukuran dilakukan menggunakan instrumen tes. Instrumen ini dirancang secara khusus untuk mengukur kemampuan mahasiswa pada materi perhitungan premi asuransi jiwa pada matakuliah matematika aktuaria. Instrumen tes dirancang dengan memperhatikan komponen konstruk dan komponen deskriptif. Komponen konstruk adalah bahwa butir soal harus memberikan indikasi terhadap jawaban yang akan yang akan didapatkan, dari tahapan yang rendah ke tahapan yang tinggi. Pada penelitian ini, butir soal akan menghasilkan jawaban yang menunjukkan rentang kualitatif tertentu yaitu dua alternatif jawaban, benar atau salah. Untuk selanjutnya disebut pola respon dikotomi. Komponen kedua adalah komponen deskriptif yaitu komponen yang menjelaskan beberapa butir pada aspek tertentu. Pada penelitian ini aspek yang di perhatikan adalah bagaimana kemampuan mahasiswa untuk memahami prinsip perhitungan premi pada asuransi jiwa. Mata kuliah matematika aktuaria diberikan kepada mahasiswa jurusan matematika Universitas Negeri Malang kemudian dianalisa menggunakan model Rasch.

Hasil tes yang berupa skor kemudian dianalisa dengan model Rasch. Model Rasch merupakan suatu teori penilaian modern yang dapat mengklasifikasikan perhitungan item dan person dalam suatu peta distribusi (Rozeha, Azami, Saidudin, 2007). Model ini merupakan bagian dari teori respon butir (Thissen, Nelson, Rosa, et al, 2001). Model rasch didasarkan pada dua prinsip. Prinsip pertama yaitu kemampuan subjek dalam hal ini kemampuan mahasiswa terhadap suatu pertanyaan yang dapat diprediksi menggunakan seperangkat faktor yang disebut dengan traits. Traits adalah suatu dimensi dari kemampuan individu yang dapat terdiri dari kemampuan verbal, kognitif serta kemampuan psikomotorik. Prinsip kedua menyatakan hubungan yang terjadi antara kemampuan subjek, dalam hal ini mahasiswa, pada suatu soal atau pertanyaan terhadap kemampuan

lain dapat digambarkan dalam suatu kurva karakteristik butir (Hambleton, Swaminathan, & Rogers, 1991. p. 7). Dalam model rasch, peserta tes dengan kemampuan (ability) yang tinggi seharusnya memiliki probability yang lebih besar untuk menjawab benar suatu soal daripada mahasiswa lainnya. Dan sebaliknya, mahasiswa memiliki peluang lebih kecil untuk menjawab benar suatu soal yang memiliki tingkat kesulitan lebih tinggi (Sumintono & Widhiarso, 2015, p. 3). Dalam pendekatan model Rasch, selain memperhatikan item juga memperhatikan aspek respond dan korelasinya (Ardiyanti, 2016. p.248-263). Keunggulan pemodelan Rasch dibanding metoda lainnya, khususnya teori tes klasik, yaitu kemampuan melakukan prediksi terhadap data yang hilang (missing data), yang didasarkan kepala pola respon yang sistematis (Aziz, R. 2015.p. 29-39)

Suatu instrument tes dikatakan baik jika mampu memberi informasi yang tepat terkait kemampuan mahasiswa atas kompetensi yang diujikan. Pada tahun 2016, susongko melakukan penelitian terkait model validasi soal tes yaitu validitas Messick yang mencakup beberapa aspek seperti aspek isi, struktural, substantif, konsekuensial dan eksternal. Validitas ini dianalisa menggunakan model Rasch yang dibantu software Winsteps. Penelitian lainnya terkait analisa teori tes dilakukan oleh Kustriyono (2004) yang menggunakan analisis kualitatif untuk mengetahui karakteristik tes yang diujikan dari sisi materi, konstruksi dan bahasa (Kustriyono. 2004. pp. 175-198).

Software Winstep merupakan alat bantu komputasi pada model Rasch untuk menganalisa skor yang dihasilkan dari instrumen tes dengan tujuan mengetahui Outfit MNSQ, Outfit ZSTD, Point Measure Correlation, Item reliability dan Alpha Cornbach. Outfit MNSQ berguna untuk melihat kesesuaian data dengan model yang digunakan. Nilai mean square yang diharapkan adalah 1 (satu). Jika Nilai mean-square pada infit lebih besar dari satu, variasi dari instrumen lebih banyak dari prediksi yang dilakukan oleh Rasch model. Jika nilai infit kurang dari 1, maka variasi pada instrumen lebih sedikit jika dibandingkan dengan prediksi yang dilakukan oleh Rasch model.

Pada penelitian ini, peneliti ingin mengetahui kualitas instrumen tes yang digunakan untuk mengetahui kemampuan mahasiswa pada materi prinsip perhitungan premi asuransi jiwa dengan pendekatan model Rasch. Kualitas ini diukur berdasarkan beberapa indicator yaitu butir soal yang fit dengan model rasch dan reliabilitas butir soal. Oleh karena itu dirancang suatu

instrumen tes kemudian ditentukan butir soal mana saja yang fit serta yang tidak fit dengan model rasch. Selain itu, dengan bantuan software Winstep akan ditentukan nilai alfa crobach untuk mengetahui reliabilitas butir soal.

2. Metode Penelitian

Penelitian ini difokuskan pada analisis instrumen tes menggunakan model Rasch. Subyek dari penelitian ini adalah mahasiswa yang mengikuti mata kuliah aktuaria di Jurusan Matematika Universitas Negeri Malang yang terdiri atas 40 mahasiswa. Terdapat 50 butir soal pada instrumen tes prestasi yang diujikan kepada mahasiswa. Tingkat kesukaran soal dimulai dari mudah, sedang dan sulit dengan perbandingan yang relative sama Soal berupa pilihan ganda dengan skor benar 1 dan skor salah 0. Sehingga data yang diperoleh merupakan data dikotomi. Hasil tes yang berupa skor dianalisa menggunakan bantuan software Winsteps. Dari output software Winsteps diperoleh beberapa parameter butir soal yang fit dengan model Rasch. Selain itu, diperoleh nilai alpha cronbach yang merupakan hasil uji reliabilitas butir secara keseluruhan. Sedangkan Outfit MNSQ, Outfit ZSTD dan nilai korelasi butir dengan soal secara keseluruhan menunjukkan batas item yang dinyatakan fit dengan model. Yaitu, apabila nilai Outfit MNSQ berada diantara 0,5 sampai dengan 1,5; nilai Outfit ZSTD berada diantara -2,0 sampai dengan 2,0; serta nilai korelasi butir dengan skor total berada diantara 0,4 sampai dengan 0,85 (Sumintono & Widhiarso, 2015, p. 12).

3. Hasil dan Pembahasan

Berdasarkan analisa data menggunakan software Winsteps, terdapat 25 Item yang fit dengan model Rasch dan 25 item lainnya tidak fit dengan model rasch. Hasil ini secara lengkap disajikan pada Tabel 1.

Tabel 1. Ringkasan Statistik

	Keterangan	Nilai
Logit	Person	0.75
	Item	0
Reliabilitas	Person Reliability	0.74
	Item Reliability	0.77
	Alpha Cronbach	0.85
Outfit MNSQ	Person	0.98
	Item	0.98
Outfit ZSTD	Person	0

Item -0.01

Pada Tabel 1 ditunjukkan nilai logit dari person atau measure sebesar 0.75 dan nilai item measure sebesar 0 yang artinya nilai person measure lebih besar dari item measure. Dapat dinyatakan bahwa abilitas mahasiswa cenderung lebih tinggi dibanding tingkat kesukaran soal Dengan kata lain, terdapat kemungkinan semua butir soal dapat dijawab dengan benar oleh mahasiswa. Sehingga mahasiswa yang memiliki abilitas tertinggi dapat menjawab soal yang paling sulit dengan benar. Sementara itu, Reliabilitas item (item reliability) bernilai 0,77, Reliabilitas Person bernilai 0.74 dan nilai Alpha Cronbach sebesar 0,85. Dari nilai tersebut dapat dinyatakan bahwa tingkat konsistensi jawaban dari siswa cukup tinggi, dan kualitas butir soal yang ada pada instrumen tes yang digunakan memiliki reliabilitas yang cukup baik yaitu 0.77. selain itu, nilai dari Alpha Cronbach yang menunjukkan interaksi antara person dan item secara keseluruhan bernilai cukup baik yaitu 0.85.

Besaran lain yang ditunjukkan pada tabel 1 adalah Nilai Outfit Mean Squared (Outfit MNSQ) sebesar 0.98 baik pada kolom person maupun item. Nilai 0.98 termasuk dalam kriteria fit yaitu terletak diantara selang $0,5 < MNSQ < 1,5$, artinya instrumen tes yang digunakan sudah sesuai dengan model untuk mengukur kompetensi mahasiswa pada materi perhitungan presmi asuransi jiwa. Nselanjutnya, diperoleh nilai Outfit Z Standardized (Outfit ZSTD) sebesar 0 untuk person dan -0,01 untuk item. Nilai 0 dan -0,01 berada diantara selang $-2,0 < ZSTD < 2,0$ yang dapat diartikan data memiliki kemungkinan nilai yang rasional. Hal ini berarti secara keseluruhan butir soal atau item telah sesuai dengan model rasch dan dapat dijadikan instrumen tes prestasi pada materi perhitungan premi asuransi jiwa.

Distribusi item soal yang dianggap misfit atau tidak fit dengan model dapat dilihat pada Tabel 2. Batas item dinyatakan fit dengan model jika memenuhi salah satu atau keduanya dari syarat berikut. Syarat pertama, nilai Outfit MNSQ terletak diantara 0,5 sampai dengan 1,5; nilai Outfit ZSTD terletak diantara -2,0 sampai dengan 2,0; dan nilai korelasi butir dengan skor total (point measure correlation) terletak diantara 0,4 sampai dengan 0,85 (Sumintono & Widhiarso, 2014, pp. 71-72).

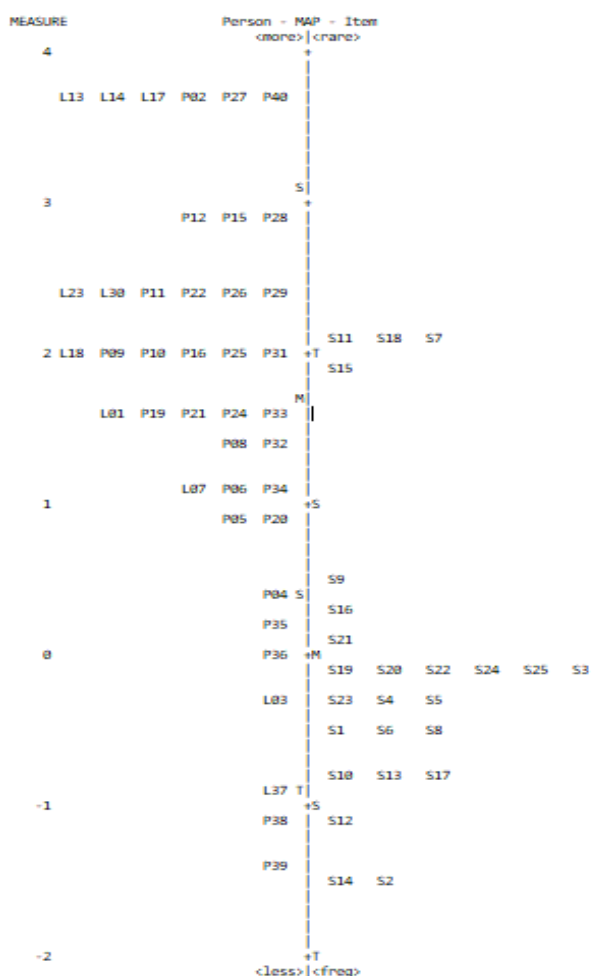
Tabel 2. Data Sebaran Soal Misfit Atau Tidak Fit Dengan Model Rasch

Item STATISTICS: MISFIT ORDER

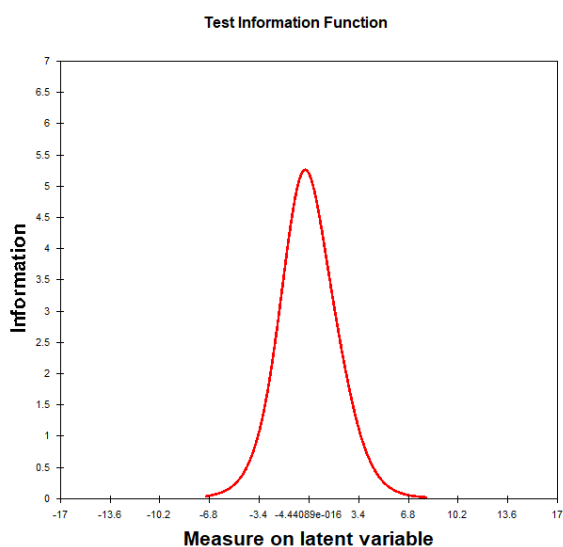
ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	MEASURE	MODEL		INFIT		OUTFIT		PTMEASUR-AL		EXACT OBS%	MATCH EXP%	Item
				S. E.	MNSQ	ZSTD	MNSQ	ZSTD	CORR.	EXP.				
1	34	40	-.51	.51	.98	.06	2.21	1.59	A	.38	.45	87.5	87.2	S1
23	33	40	-.27	.48	.94	-.11	1.95	1.49	B	.42	.46	87.5	85.6	S23
14	37	40	-1.49	.66	1.46	1.05	1.74	.92	C	.05	.36	92.5	92.4	S14
3	32	40	-.05	.46	1.45	1.53	1.11	.39	D	.26	.47	75.0	84.0	S3
6	34	40	-.51	.51	1.43	1.29	1.41	.77	E	.20	.45	82.5	87.2	S6
5	33	40	-.27	.48	1.42	1.37	1.04	.26	F	.27	.46	77.5	85.6	S5
4	33	40	-.27	.48	1.41	1.32	1.00	.19	G	.29	.46	77.5	85.6	S4
12	36	40	-1.10	.59	1.39	1.01	1.15	.46	H	.20	.40	85.0	90.3	S12
18	18	40	2.06	.36	1.27	1.79	1.37	1.28	I	.29	.48	57.5	70.4	S18
2	37	40	-1.49	.66	1.25	.66	.91	.29	J	.25	.36	92.5	92.4	S2
7	18	40	2.06	.36	1.20	1.39	1.17	.69	K	.35	.48	62.5	70.4	S7
21	31	40	.15	.44	.96	-.09	1.10	.37	L	.48	.48	85.0	82.4	S21
22	32	40	-.05	.46	.92	-.21	.89	-.04	M	.51	.47	90.0	84.0	S22
15	19	40	1.93	.36	.87	-.89	.74	-.98	l	.58	.48	70.0	70.2	S15
11	18	40	2.06	.36	.85	-1.04	.72	-1.00	k	.59	.48	72.5	70.4	S11
19	32	40	-.05	.46	.83	-.59	.79	-.28	j	.56	.47	90.0	84.0	S19
20	32	40	-.05	.46	.83	-.56	.67	-.56	i	.58	.47	85.0	84.0	S20
24	32	40	-.05	.46	.83	-.56	.67	-.56	h	.58	.47	85.0	84.0	S24
25	32	40	-.05	.46	.83	-.56	.67	-.56	g	.58	.47	85.0	84.0	S25
13	35	40	-.78	.54	.76	-.65	.56	-.43	f	.56	.42	95.0	88.8	S13
17	35	40	-.78	.54	.76	-.65	.56	-.43	e	.56	.42	95.0	88.8	S17
16	30	40	.33	.42	.70	-1.34	.68	-.72	d	.66	.49	90.0	81.0	S16
8	34	40	-.51	.51	.68	-1.01	.54	-.62	c	.62	.45	92.5	87.2	S8
9	29	40	.50	.41	.58	-2.19	.42	-1.87	b	.76	.49	90.0	79.5	S9
10	35	40	-.78	.54	.57	-1.33	.30	-1.04	a	.67	.42	95.0	88.8	S10
MEAN	30.8	40.0	.00	.48	1.01	.0	.98	.0				83.9	83.5	
P.SD	5.8	.0	1.01	.08	.29	1.1	.46	.8				9.8	6.5	

Berdasarkan hasil analisis instrumen tes prestasi dengan menggunakan bantuan program winsteps version 4.4.5 yang ada pada tabel 2 diperoleh butir soal yang misfit 25 butir, dan soal fit sebanyak 25 butir, sehingga diperoleh instrumen final sebanyak 25 butir.

Nilai tingkat kemampuan mahasiswa dalam mengerjakan soal diperlihatkan dari output winsteps yaitu peta wright. Diperoleh data mahasiswa dengan kode L13, L14, L17, P02, P27 dan P40 memiliki kemampuan atau abilitas paling tinggi sedangkan mahasiswa P39 memiliki kemampuan paling rendah.



Gambar 1. Person Measure



Gambar 2. Test Information Function

Gambar 2 menunjukkan tentang informasi pengukuran yang didapatkan dari instrumen tes prestasi mahasiswa pada mata kuliah matematika aktuaria. Sumbu-x menunjukkan tingkat kemampuan siswa dalam mengerjakan tes yang diberikan, sedangkan sumbu-y menunjukkan nilai dari fungsi informasi. Berdasarkan grafik tersebut,

pada level kemampuan sedang, informasi yang didapatkan oleh pengukuran sangat tinggi. Sehingga pengembangan instrumen tes prestasi mahasiswa pada mata kuliah matematika aktuaria cocok atau optimal jika digunakan pada siswa dengan kemampuan sedang.

4. Kesimpulan

Instrumen tes yang digunakan untuk mata kuliah matematika aktuaria pada materi perhitungan premi asuransi jiwa fit dengan model rasch. Hal ini ditunjukkan dengan skor item (item reliability) sebesar 0.77, reliabilitas person (person reliability) sebesar 0.74, dan nilai alpha cronbach sebesar 0.85 sedangkan nilai Outfit Mean Square Statistic (Outfit MNSQ) sebesar 0.98 pada kolom person maupun item. Nilai Outfit Z Standard (Outfit ZSTD) sebesar 0 pada tabel person dan sebesar -0.01 untuk tabel item. Sedangkan umlah butir soal yang fit sebanyak 25 sedangkan yang tidak fit sebanyak 25 butir soal.

Daftar Pustaka

Ardiyanti, D. (2016). Aplikasi Model Rasch pada Pengembangan Skala Efikasi Diri dalam Pengambilan Keputusan Karir Siswa. *Jurnal Psikologi*, 43(3), 248-263.

Aziz, R. (2015). Aplikasi model Rasch dalam pengujian alat ukur kesehatan mental di tempat kerja. *Psikoislamika: Jurnal Psikologi dan Psikologi Islam*, 12(2), 29-39.

Dunn, L., Morgan, C., O'Reilly, M., & Parry, S. (2003). *The student assessment handbook: New directions in traditional and online assessment*. Routledge.

Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of item response theory*. London: SAGE Publications. dalam *Rangka Pengembangan Bank Soal*. *Jurnal Penelitian dan Evaluasi Pendidikan*, 6(2).

Kustriyono. (2004). Penyusunan Perangkat Soal Ujian Akhir Mata Pelajaran Sains Biologi Dalam Rangka Pengembangan Bank Soal. *Jurnal Penelitian dan Evaluasi Pendidikan* 2 (VI): 175-198.

Sumintono, B., & Widhiarso, W. (2014). Aplikasi model Rasch untuk penelitian ilmu-ilmu sosial (edisi revisi). Bandung: Trim Komunikata Publishing House.

Rozeha, A. R., Azami, Z. & Mohd Saidudin, M. (2007). Application of Rasch Measurement in Evaluation of Learning Outcomes: A Case Study in Electrical Engineering. *Regional Conference on Engineering Mathematics, Mechanics, Manufacturing & Architecture 2007 (EM3ARC)*

- Sumintono, B., & Widhiarso, W. (2015). Aplikasi pemodelan rasch pada assessment pendidikan. Trim Komunikata.
- Swaminathan, H., Hambleton, R. K., & Rogers, H. J. (2007). Assessing the fit of item response models. *Handbook of Statistics*, 26, 683-718.
- Thissen, D., Nelson, L., Rosa, K., et al. 2001. Item Response Theory for Items Scored in More than Two Categories dalam D. Thissen & H. Wainer, *Test Scoring* (pp. 141-184). New Jersey: Lawrence Erlbaum Associates Publishers.