

ANALISIS BUTIR SOAL MATEMATIKA PADA UKA PLPG LPTK FAKULTAS TARBIYAH IAIN ANTASARI BANJARMASIN

Oleh Rahmawati

Abstrak

Penelitian ini bertujuan untuk mengetahui kualitas butir soal matematika pada UKA PLPG LPTK Fakultas Tarbiyah IAIN Antasari Banjarmasin pada tahun 2012. Metode yang digunakan dalam penelitian ini deskriptif kuantitatif. Hasil analisis menunjukkan bahwa kualitas butir soal dari tingkat validitas soal, yang valid sebesar 66,7 % dan tidak valid sebesar 33,3 %. Dari tingkat reliabilitas soal dengan taraf signifikansi 5% soal dinyatakan tidak reliabel. Berdasarkan tingkat kesukaran soal 83,33% berada pada kategori mudah dan 16,67% berada pada kategori sedang. Daya beda soal menunjukkan 66,6% soal dikategorikan baik, 16,7% dengan kategori cukup, dan 16,7% dengan kategori tidak baik. Efektivitas distraktor dari 5 soal yang diujikan, distraktor yang berfungsi hanya 1 butir soal, sedangkan yang tidak berfungsi sebanyak 4 butir soal.

Kata kunci: Butir soal matematika, dan UKA

A. Pendahuluan

Sertifikasi guru sebagai upaya meningkatkan profesionalitas guru di Indonesia, diselenggarakan berdasarkan landasan hukum Peraturan Menteri Pendidikan Nasional Nomor 16 Tahun 2005 tentang Standar Kualifikasi dan Kompetensi Pendidik dan Peraturan Menteri Pendidikan dan Kebudayaan Republik Indonesia Nomor 5 Tahun 2012 tentang Sertifikasi bagi Guru Dalam Jabatan. Pendidikan dan Latihan Profesi Guru (PLPG) bertujuan untuk meningkatkan kompetensi, profesionalisme, dan menentukan kelulusan guru peserta sertifikasi.

Pada akhir kegiatan PLPG dilakukan uji kompetensi dengan mengacu pada rambu-rambu pelaksanaan PLPG. Uji kompetensi meliputi uji tulis dan uji kinerja (ujian praktik). Ujian tulis bertujuan untuk mengungkap kompetensi profesional dan pedagogik. Terkait butir-butir soal yang diujikan haruslah representatif mengukur kompetensi guru sesuai dengan karakteristik mata pelajaran dari peserta uji kompetensi. Salah satu cara untuk mengetahui

butir soal yang representatif dan memiliki kualitas baik maka perlu dilakukan analisis butir soal.

Kenyataan di lapangan menunjukkan bahwa selama ini kegiatan analisis butir-butir soal jarang dilakukan. Itulah sebabnya materi, konstruksi soal, bahasa, validitas, reliabilitas, dan analisis butir soal yang terdiri dari tingkat kesukaran, daya pembeda, dan distraktor soal sering dikatakan rendah. Lebih tepatnya kualitasnya tidak diketahui secara pasti, termasuk soal matematika pada UKA (ujian kompetensi akhir) PLPG LPTK Fakultas Tarbiyah IAIN Antasari Banjarmasin tahun 2012, sehingga belum dapat dipastikan apakah soal tersebut sudah memenuhi syarat-syarat tes yang baik atau belum.

Terkait dengan tidak dilakukannya uji validasi terhadap butir soal matematika pada UKA PLPG LPTK Fakultas Tarbiyah, menyebabkan karakteristik dan kualitas tes belum diketahui. Faktor kualitas tes yang belum diketahui, akan berpengaruh terhadap kemampuan testee dalam mengerjakan tes. Kelemahan ini akan berdampak pada sulitnya menentukan kemampuan testee yang sebenarnya.

Salah satu cara mengantisipasi keadaan seperti di atas adalah dengan jalan melakukan sebuah analisis terhadap butir soal matematika pada UKA PLPG LPTK Fakultas Tarbiyah IAIN Antasari Banjarmasin tahun 2012. Apabila analisis item soal dilakukan secara cermat, maka dapat diketahui apakah butir-butir soal tersebut sudah dapat menjalankan fungsinya sebagai alat pengukur hasil UKA yang memadai atau belum.

Analisis setiap soal matematika pada UKA PLPG LPTK Fakultas Tarbiyah IAIN Antasari Banjarmasin tahun 2012 itu dilakukan dengan harapan akan menghasilkan berbagai informasi yang berharga, guna melakukan perbaikan, pembenahan, dan penyempurnaan terhadap butir-butir item yang telah dikeluarkan dalam tes UKA. Sehingga pada masa-masa yang akan datang tes yang disusun atau dirancang oleh tester (dosen) itu betul-

betul dapat menjalankan fungsinya sebagai alat pengukur hasil belajar yang memiliki kualitas yang tinggi.¹

Matematika merupakan salah satu materi yang diujikan pada UKA yang mempunyai peranan penting dalam upaya mewujudkan tujuan sebuah pendidikan. Matematika merupakan sumber dari ilmu-ilmu lain, banyak ilmu-ilmu lain yang penemuan dan perkembangannya bergantung dari matematika. sehingga sangatlah perlu diperhatikan pengujian validasi terhadap soal tersebut saat pembuatannya.

Berdasarkan permasalahan diatas, maka peneliti tertarik untuk mengetahui bagaimana kualitas Butir soal Matematika pada UKA (ujian kompetensi akhir) PLPG LPTK Fakultas Tarbiyah IAIN Antasari Banjarmasin Tahun 2012 dari segi: validitas, reliabilitas, tingkat kesukaran, daya pembeda, dan efektivitas pengecohnya.

B. TINJAUAN TEORITIS

1. Kualitas Tes

Menurut Wayan Nurkencana dan Sunartana mengemukakan kualitas atau baik buruknya suatu tes dapat ditinjau dari beberapa segi yaitu: validitas, reliabilitas, tingkat kesukaran, daya pembeda, dan distraktor (pengecoh).

a. Validitas

Pengertian umum untuk validitas item adalah sebuah item dikatakan valid apabila mempunyai dukungan yang besar terhadap skor total. Skor pada item menyebabkan skor total menjadi tinggi atau rendah. Dengan kata lain dapat dikemukakan di sini bahwa sebuah item memiliki validitas yang tinggi jika skor pada item mempunyai kesejajaran dengan skor total. Kesejajaran ini dapat diartikan dengan

¹ Anas Sudiyono, *Pengantar Evaluasi Pendidikan*, (Jakarta: PT RajaGrafindo Persada, 2003), h.370

korelasi sehingga untuk mengetahui validitas item digunakan rumus korelasi.²

b. Reliabilitas

Reliabilitas sering diartikan keterandalan. Artinya suatu tes memiliki keterandalan jika tes tersebut dipakai mengukur berulang-ulang hasilnya sama.³ Ada beberapa cara yang dapat digunakan untuk mengadakan uji reliabilitas tes, namun karena tes dalam penelitian ini berbentuk pilihan ganda maka untuk mengetahui koefisien reliabilitas tes soal bentuk pilihan ganda digunakan rumus Kuder Richadson 20 (KR-20).

c. Tingkat Kesukaran

Tingkat kesukaran soal adalah peluang untuk menjawab benar suatu soal pada tingkat kemampuan tertentu yang biasanya dinyatakan dalam bentuk indeks. Soal yang baik adalah soal yang tidak terlalu mudah atau tidak terlalu sukar.⁴ Soal yang terlalu mudah tidak merangsang seseorang untuk mempertinggi usaha memecahkannya. Sebaliknya, soal yang terlalu sukar akan menyebabkan seseorang menjadi putus asa dan tidak mempunyai semangat untuk mencoba lagi karena di luar jangkauannya.

Indeks tingkat kesukaran pada umumnya dinyatakan dalam bentuk proporsi yang besarnya berkisar antara 0,00 – 1,00. Semakin besar indeks tingkat kesukaran yang diperoleh dari hasil hitungan, berarti semakin mudah soal itu. Soal dengan indeks kesukaran 0,00 menunjukkan bahwa soal itu terlalu sukar, karena ini menunjukkan bahwa tidak ada testee yang dapat menjawab soal itu dengan betul. Sebaliknya jika indeks kesukarannya 1,00 maka menunjukkan bahwa soal tersebut terlalu mudah, karena hal ini berarti seluruh testee dapat menjawab soal itu dengan benar.

d. Daya Pembeda

² Suharsimi Arikunto, *op. cit.*, h. 76.

³ M. Chabib Thoah, *op. cit.*, h. 118.

⁴ Daryanto, *Evaluasi Pendidikan*, (Jakarta: Rineka Cipta, 2008), Cet. ke-5, h.179

Daya pembeda adalah kemampuan suatu soal untuk membedakan antara seseorang yang berkemampuan tinggi dengan seseorang yang berkemampuan rendah.⁵ Daya pembeda suatu soal berfungsi untuk menentukan dapat tidaknya suatu soal membedakan kelompok dalam aspek yang diukur sesuai dengan perbedaaan yang ada pada kelompok itu. Manfaat daya pembeda butir soal adalah sebagai berikut :

- 1) Untuk meningkatkan mutu setiap butir soal melalui data empiriknya. Berdasarkan indeks daya pembeda setiap butir soal dapat diketahui apakah butir soal itu baik, direvisi atau ditolak.
- 2) Untuk mengetahui seberapa jauh setiap butir soal dapat mendeteksi atau membedakan kemampuan seseorang, yaitu seseorang yang telah memahami atau yang belum memahami materi yang diajarkan guru.

Angka yang menunjukkan besarnya daya pembeda disebut indeks diskriminasi (D). Seperti halnya indeks kesukaran, indeks diskriminasi (daya pembeda) berkisar antara 0,00-1,00.

e. Distraktor (Pengecoh)

Pengecoh berfungsi sebagai pengidentifikasi peserta tes yang berkemampuan tinggi. Pengecoh dikatakan berfungsi efektif apabila banyak dipilih oleh peserta tes yang berasal dari kelompok bawah. Sebaliknya apabila pengecoh itu banyak dipilih oleh peserta tes yang berasal dari kelompok atas, maka pengecoh itu tidak berfungsi sebagaimana mestinya.⁶

Pengecoh yang tidak dipilih sama sekali oleh testee menyatakan bahwa pengecoh itu jelek, terlalu mencolok, menyesatkan. Sebaliknya, sebuah distraktor (pengecoh) dapat dikatakan berfungsi dengan baik apabila distraktor tersebut mempunyai daya tarik yang besar bagi pengikut-pengikut tes yang kurang memahami konsep atau kurang

⁵ Daryanto, *op. cit.*, h. 179

⁶ Nur Khoiriyah, *op. cit.*, h. 33.

menguasai bahan.⁷ Suatu distraktor dapat dikatakan baik jika paling sedikit oleh 5% peserta tes. Apabila pengecoh dipilih secara merata, maka termasuk pengecoh yang sangat baik. Dari segi omit, sebuah item dikatakan baik jika omitnya tidak lebih dari 10% pengikut tes.⁸

C. METODE PENELITIAN

Jenis penelitian yang digunakan adalah penelitian lapangan (*field research*) dengan pendekatan deskriptif. Dalam penelitian ini peneliti ingin mendeskripsikan mengenai kualitas butir soal matematika pada UKA PLPG LPTK Fakultas Tarbiyah IAIN Antasari Banjarmasin tahun 2012 dari segi: validitas, reliabilitas, tingkat kesukaran, daya pembeda, dan efektivitas distraktor (pengecoh) soal.

1. Subjek dan Objek Penelitian

a. Subjek penelitian

Subjek dari penelitian ini adalah tes UKA peserta PLPG LPTK Fakultas Tarbiyah tahun 2012 yang berjumlah sebanyak 29 orang.

b. Objek penelitian

Objek penelitian ini adalah kualitas butir soal matematika pada UKA PLPG LPTK Fakultas Tarbiyah IAIN Antasari Banjarmasin tahun 2012 dari segi: validitas, reliabilitas, tingkat kesukaran, daya pembeda, dan efektivitas pengecohnya.

2. Data dan Sumber Data

a. Data

Data yang digali dari penelitian ini yaitu data hasil tes matematika pada UKA PLPG berupa lembar jawaban soal.

⁷ Suharsimi Arikunto, *op. cit.*, h. 220

⁸ *Ibid.*, h. 221

b. Sumber Data

Dokumenter yaitu arsip-arsip dari LPTK Fakultas Tarbiyah berupa naskah soal, dan kunci jawaban soal.

3. Teknik Pengumpulan Data

Data dalam penelitian ini diperoleh melalui hasil tes matematika pada UKA PLPG LPTK Fakultas Tarbiyah IAIN Antasari Banjarmasin tahun 2012

4. Teknik Pengolahan Data

- a. *Editing* (mempelajari kembali), yaitu meneliti, memeriksa, dan mengecek kembali data-data yang sudah terkumpul untuk memastikan kelengkapannya.
- b. *Koding*, yaitu mengklasifikasikan data yang diperoleh menurut macamnya dengan memberikan kode pada setiap data yang sudah diperoleh untuk mempermudah peneliti di dalam melakukan penelitian ini.
- c. *Tabulating*, yaitu mengolah data sedemikian rupa ke dalam bentuk tabel, sesuai tabel yang diperlukan.

5. Teknik Analisis Data

Teknik analisis data dilakukan dengan menggunakan program ANATES.

a. Menentukan Validitas

Rumus yang digunakan dalam menentukan validitas soal adalah rumus korelasi *Product Moment* dengan angka kasar sebagai berikut:

$$r_{xy} = \frac{N \sum XY - (\sum X)(\sum Y)}{\sqrt{\{N \sum X^2 - (\sum X)^2\} \{N \sum Y^2 - (\sum Y)^2\}}}$$

Dengan : r_{xy} = koefisien korelasi product moment
N = jumlah objek
X = skor butir soal

$$Y = \text{skor total}^9$$

Untuk memberikan interpretasi terhadap r_{xy} , digunakan tabel r product moment. Jika $r_{xy} \geq r_{tabel}$, dengan taraf signifikansi 5% maka butir soal tersebut dikatakan valid.

b. Menentukan reliabilitas

Rumus yang digunakan untuk menentukan reliabilitas tes adalah dengan menggunakan rumus *Kuder Richardson-20 (KR-20)*. Metode KR-20 ini berguna untuk mengetahui reliabilitas dari seluruh tes untuk item pertanyaan atau pernyataan yang menggunakan jawaban benar (YA) atau salah (TIDAK). Bila benar bernilai = 1 dan jika salah bernilai = 0. Rumus yang digunakan adalah sebagai berikut:

$$r_{11} = \left(\frac{k}{k-1} \right) \left(\frac{s^2 - \sum pq}{s^2} \right)$$

Dengan:

- r_{11} = Koefisien reliabilitas internal seluruh item
- p = Proporsi subjek yang menjawab item dengan benar
- q = Proporsi subjek yang menjawab item dengan salah
($q = 1 - p$)
- $\sum pq$ = jumlah hasil perkalian p dengan q
- k = banyaknya item
- s = standar deviasi dari tes¹⁰

c. Menentukan tingkat kesukaran (*Difficulty Level*)

Untuk mengukur tingkat kesukaran digunakan rumus:

$$P = \frac{B}{JS}$$

Dimana:

P = Proporsi angka indeks kesukaran item.

B = Banyaknya testee yang dapat menjawab soal itu dengan betul.

⁹ Suharsimi Arikunto, *Prosedur Penelitian Suatu Pendekatan Prektek*, (Jakarta: Rineka Cipta, 2002), h. 171

¹⁰ Riduwan, *Belajar Mudah Penelitian Untuk Guru, Karyawan, dan Peneliti Pemula*, (Bandung: Alfabeta, 2004), h. 108

JS = jumlah testee yang mengikuti tes hasil belajar.¹¹

Mengenai bagaimana cara memberikan penafsiran (interpretasi) terhadap angka indeks kesukaran item, Witherington dalam bukunya berjudul *Psychological Evaluation* mengemukakan sebagai berikut:

Tabel 1.1 Interpretasi Indeks Kesukaran Item¹²

Besarnya P	Interpretasi
< 0,25	Sukar
0,25 – 0,75	sedang
> 0,75	mudah

d. Menentukan daya pembeda (*Discriminating power*)

Rumus untuk menentukan indeks diskriminasi adalah:

$$D = \frac{B_A}{J_A} - \frac{B_B}{J_B} = P_A - P_B$$

Dimana:

J = Jumlah peserta tes.

J_A = Banyak peserta kelompok atas.

J_B = Banyak peserta kelompok bawah.

B_A = Banyak peserta kelompok atas yang menjawab soal itu dengan benar.

B_B = Banyak peserta kelompok bawah yang menjawab soal itu dengan benar.

P_A = proporsi peserta kelompok atas yang menjawab benar.

P_B = proporsi peserta kelompok bawah yang menjawab benar.¹³

Menurut Fernandes (1984) kategori indeks daya pembeda butir soal adalah sebagai berikut:

¹¹ Anas Sudiyono, *op. cit.*, h. 372

¹² *Ibid.*

¹³ *Ibid.*, h. 186

Tabel 1.2. Interpretasi Daya Pembeda Item¹⁴

Besarnya Indeks Daya Pembeda (D)	Klasifikasi
$D \geq 0,40$ $0,30 \leq D \leq 0,39$ $0,20 \leq D \leq 0,29$ $D \leq 0,19$ D bertanda negative	Sangat baik (<i>excellent</i>) Baik (<i>good</i>) Cukup (<i>satisfactory</i>) Tidak baik (<i>poor</i>) Sangat tidak baik

Akhirnya sebagai tindak lanjut atas hasil penganalisaan mengenai daya pembeda butir soal UKA tersebut adalah:

1. Butir soal yang telah memiliki daya pembeda item yang baik (*satisfactory, good, dan excellent*) hendaknya dimasukkan dalam bank soal dan bisa digunakan kembali pada tes yang akan datang.
2. Butir soal yang daya pembedanya masih rendah (*poor*) ada 2, yaitu: diperbaiki atau didrop.
3. Khusus butir soal yang angka indeksinya bertanda negatif, itu menandakan daya pembeda yang sangat tidak baik, sebaiknya pada tes hasil belajar yang akan datang tidak dikeluarkan kembali.¹⁵
- e. Menentukan keefektifan pengecoh (*distractor*)

Dalam penelitian ini penulis dalam menentukan efektivitas pengecoh berpatokan pada kriteria jumlah kelompok atas dan kelompok bawah yang memilih option pengecoh sebagai berikut:

- 1) Pengecoh berfungsi efektif, jika $N_A < N_B$
- 2) Pengecoh tidak berfungsi efektif, jika $N_A = N_B$
- 3) Pengecoh menyesatkan, jika $N_A > N_B$

Keterangan:

N_A = Jumlah peserta tes kelompok atas yang memilih pengecoh

N_B = Jumlah peserta tes kelompok bawah yang memilih pengecoh

¹⁴ Yanugilang, "Analisis Perbaikan Butir Soal", <http://yanugilang.blogspot.com/2010/02/op.html/top>

¹⁵ Hendriyansyah, "Cara Menganalisis Butir Soal", <http://hendriyansyah.blogspot.com/2010/04/op.html/top>

D. HASIL PENELITIAN DAN PEMBAHASAN

1. Validitas

Perhitungan validitas menggunakan rumus Korelasi Product Moment. Diketahui jumlah peserta tes (N) = 29 orang dengan taraf signifikansi 5%. Jika $r_{xy} > r_{tabel}$ maka soal UKA dikatakan valid, sebaliknya jika $r_{xy} \leq r_{tabel}$ maka soal dikatakan tidak valid. Adapun hasil uji validitas butir soal dapat tabel ditentukan persentasi validitas sebagai berikut:

Tabel 2.2. Persentasi Validitas Soal

Kelompok	Frekuensi	Persentasi (%)
Valid	4	66,7
Tidak valid	2	33,3
Jumlah	6	100

Dari 6 butir soal hasil tes matematika pada UKA sertifikasi guru PLPG Fakultas Tarbiyah, terdapat 66,7% (4 buah) soal yang dinyatakan valid dan 33,3% (2 buah) soal yang dinyatakan tidak valid.

2. Reliabilitas

Berdasarkan perhitungan reliabilitas dengan menggunakan rumus KR-20 dapat diketahui bahwa indeks reliabilitas soal matematika Pada UKA PLPG LPTK Fakultas Tarbiyah IAIN Antasari Banjarmasin Tahun 2012 adalah 0,21. Telah diketahui sebelumnya bahwa untuk taraf 5% dan $N = 29$ $r_{tabel} = 0,34$. Karena $0,21 < 0,34$ yang artinya $r_{11} < r_{tabel}$, maka perangkat soal tergolong tidak reliabel.

3. Tingkat Kesukaran

Berdasarkan perhitungan yang diperoleh hasil tingkat kesukaran (P) soal matematika Pada UKA PLPG LPTK Fakultas Tarbiyah IAIN Antasari Banjarmasin Tahun 2012 dapat diketahui bahwa rata-rata soal diatas berada pada kategori mudah yaitu 0,80. Untuk persentase tingkat kesukarannya dapat dilihat pada tabel berikut:

Tabel 2.4 Persentasi Tingkat Kesukaran Soal

Tingkat Kesukaran (P)	Kategori	Frekuensi	Persentasi (%)
< 0,25	Sukar	0	0
0,25 - 0,75	Sedang	1	16,67
> 0,75	Mudah	5	83,33
Jumlah		6	100

Berdasarkan temuan di atas, dapat dinyatakan bahwa tingkat kesukaran soal matematika pada UKA PLPG LPTK Fakultas Tarbiyah IAIN Antasari Banjarmasin Tahun 2012 dianggap belum cukup baik, sehingga perlu dilakukan revisi terutama butir soal no, 1, 2, 3, 4 dan 6 yang memiliki tingkat kesukaran berkategori mudah.

4. Daya Pembeda Soal

Berdasarkan tabel di atas dapat ditentukan persentasi daya pembeda soal, yaitu sebagai berikut:

Tabel 2.6. Persentasi Daya Pembeda

Daya Pembeda (D)	Kategori	Frekuensi	Persentasi (%)
$D \geq 0,40$	Sangat baik	0	0
$0,30 \leq D \leq 0,39$	Baik	4	66,6
$0,20 \leq D \leq 0,29$	Cukup	1	16,7
$D \leq 0,19$	Tidak baik	1	16,7
Bertanda negative	Sangat tidak baik	0	0
Jumlah		6	100

Dari tabel 2.6 di atas dapat diketahui bahwa dari 6 soal matematika pada UKA PLPG LPTK Fakultas Tarbiyah IAIN Antasari Banjarmasin Tahun 2012 yang dianalisis, terdapat 66,6% soal dengan kategori daya pembeda baik, 16,7% dengan kategori cukup, dan 16,7% dengan kategori tidak baik yang sebaiknya dibuang saja.

5. Efektivitas Distraktor (Pengecoh) Soal

Berdasarkan hasil analisis UKA dapat ditentukan efektivitas distraktor soal sebagai berikut:

Tabel 2.7. Efektivitas Distraktor Soal

No.	Alternatif Jawaban			
	A	B	C	D
1	2	3	4	5
1	B	TB	B	K
2	K	B	B	B
3	TB	B	B	K
4	TB	K	TB	B
5	TB	TB	K	B
6	B	K	B	TB

Keterangan : B = Pengecoh berfungsi

TB = Pengecoh tidak berfungsi

M = Pengecoh menyesatkan

K = Kunci jawaban

Berdasarkan tabel 2.7 di atas dapat diketahui pengecoh yang tidak berfungsi berjumlah sebanyak 4 buah, yaitu pada soal nomor 1 (B), 3 (A), 4(A dan C), 5 (A dan B), dan 6 (D).

E. Kesimpulan

Berdasarkan hasil analisis, kualitas butir soal matematika pada UKA PLPG LPTK Fakultas Tarbiyah IAIN Antasari Banjarmasin Tahun 2012 dapat digambarkan sebagai berikut:

1. Dari segi validitas, soal yang diklasifikasikan valid sebesar 66,7%. sedangkan soal yang diklasifikasikan tidak valid sebesar 33,3%.
2. Dari segi reliabilitas soal, yaitu 0,21 dengan taraf signifikansi 5% dan $r_{tabel} = 0,34$, karena $0,21 < 0,34$ yaitu $r_{hitung} < r_{tabel}$ maka perangkat soal dinyatakan tidak reliabel.
3. Dari segi tingkat kesukaran, rata-rata soal adalah 0,80 berada pada kategori mudah.
4. Dari segi daya pembeda, rata-rata soal adalah 0,28 berada pada kategori cukup.
5. Dari segi efektivitas pengecoh, pengecoh yang berfungsi sebanyak 1 buah, dan yang tidak berfungsi sebanyak 5 buah.

DAFTAR PUSTAKA

- Arikunto, Suharsimi, *Dasar-dasar Evaluasi Pendidikan*, Jakarta: Bumi Aksara, 1997.
- Azwar, Saifuddin, *Tes Prestasi Fungsi dan Pengembangan Pengukuran Prestasi Belajar*, Yogyakarta: Pustaka belajar, 2000.
- Daryanto, *Evaluasi Pendidikan*, Jakarta: Rineka Cipta, 2008.
- Hadi, Amirul dan Haryanto, *Metodologi Penelitian Pendidikan*, Bandung: Pustaka Setia, 2005.
- Iskandar, *Metodologi Penelitian Pendidikan dan Sosial (Kuantitatif dan Kualitatif)*, Jakarta: Gaung Persada Press (GP Press), 2009.
- Nurkancana, Wayan dan P.P.N. Sunartana, *Evaluasi Pendidikan*, Surabaya: Usaha Nasional, 1986.
- Purwanto, M. Ngalm, *Prinsip-prinsip dan Teknik Evaluasi Pengajaran*, Bandung: PT Remaja Rosdakarya, 2002.
- Riduwan, *Belajar Mudah Penelitian Untuk Guru, Karyawan, dan Peneliti Pemula*, Bandung: Alfabeta, 2004.
- Slameto, *Evaluasi pendidikan*, Jakarta: Bina Aksara, 1998.
- Sudjana, Nana, *Penilaian Hasil Proses Belajar Mengajar*, Bandung: PT Remaja Rosdakarya, 2008.
- Sudiyono, Anas, *Pengantar Evaluasi Pendidikan*, Jakarta: PT RajaGrafindo Persada, 2003.
- Sukardi, M., *Evaluasi Pendidikan Prinsip dan Operasionalnya*, Jakarta: Bumi Aksara, 2011.
- Thoha, M. Chabib, *Teknik Evaluasi Pendidikan*, Jakarta: PT. RajaGrafindo Persada, 2003.