

ABSTRAK

Vonny Muliawati, Analisis Butir Soal Ulangan Blok Semester 1 Bahasa Jerman di SMA Kelas XII. Skripsi, Jakarta: FBS UNJ, Juli 2010

Penelitian ini bertujuan untuk mengetahui kualitas soal Ulangan Bahasa Jerman di SMA Kelas XII Tahun ajaran 2008/2009. Metode kuantitatif deskriptif digunakan dalam penelitian ini. Populasi sampling penelitian ini adalah 50 butir soal pada Ulangan Blok Semester 1 Bahasa Jerman Kelas XII di SMAN 99 Tahun ajaran 2008/2009, dengan menggunakan teknik sampling purposif.

Hasil penelitian menyatakan bahwa 50 butir soal Ulangan Blok Semester 1 Bahasa Jerman Kelas XII di SMAN 99 Jakarta dinyatakan bahwa 41 butir soal (82%) valid dan 50 butir soal (100%) reliabel. Sedangkan hasil analisis yang mencakup tingkat kesulitan, daya pembeda, pola jawaban, pada butir soal yang baik adalah 13 butir soal (26%). Butir-butir soal tersebut baik, karena memiliki tingkat kesulitan yang sedang, dapat membedakan siswa pintar dan siswa bodoh. Butir soal yang jelek pada penelitian ini sebanyak 37 butir soal (74%). Hal ini karena, butir soal memiliki tingkat kesulitan yang tidak dapat membedakan siswa pintar dan siswa bodoh.

Kesimpulan penelitian ini adalah butir soal Ulangan Blok Semester 1 Bahasa Jerman Kelas XII di SMAN 99 Jakarta Tahun Pelajaran 2008/2009 mempunyai validitas dan reliabilitas soal yang berkualitas baik. Meskipun ada butir soal yang berkualitas kurang baik.

Oleh karena itu, disarankan guru mengevaluasi butir-butir soal ulangan lebih dahulu, agar dapat mengetahui butir soal yang perlu direvisi atau yang harus diganti. Melalui penelitian ini diharapkan guru dapat terbantu dalam menganalisis butir soal dalam Ulangan Blok Bahasa Jerman .

ZUSAMMENFASSUNG

VONNY MULIAWATI. Die Itemanalyse des Deutschunterrichtstests im 1. Semester der 12. Klasse Oberschule in Jakarta. Eine wissenschaftliche Abschlussarbeit zur Erlangung des Titels „Sarjana Pendidikan“. Jakarta: Deutschabteilung der Sprach- und Kunstfakultät der Universität Negeri Jakarta, Juli 2010

I. EINFÜHRUNG

1.1 Hintergrund der Forschung

Der Deutschunterricht an der SMA wird sich an dem KTSP Lernplan orientiert. Eine Evaluation wird gebraucht, um das Ergebnis vom Lernprozess zu wissen. Mit der Evaluation werden die Lernkompetenz und das Lernergebnis der Schüler erfahren können. Um das Ziel der Evaluation zu erreichen, brauchen wir einen Test. Dazu soll der Test eine gute Qualität haben, weil das Evaluationsergebnis den Lehrern die *Feedbacks* gibt, damit sie den Lernprozess erhöhen können.

Aber in der Tatsache gibt der Test den Lehrern keine gute *Feedbacks*, weil die Schüler noch schlechte Noten in dem Test bekommen. Es besteht die Möglichkeit, dass der Test den Erfolg des Lernprozesses nicht messen können. Deshalb soll man eine Analyse für den Test machen, um die Qualität zu wissen. In der Itemanalyse wird jedes *Item* des Tests analysiert. Von jedem *Item* des

Tests werden Schwierigkeitsgrad, Trennschärfe, Distraktoren, Validität, und Reliabilität analysiert.

1.2 Identifizierung der Frage

Aufgrund des Hintergrundes der Forschung werden folgende Fragen identifiziert, nämlich:

1. Wie ist der Erfolg des Tests, der von den Lehrern die Lernleistungsmessung der Schüler verwendet wird?
2. Stimmt das Material des Deutschunterrichtstests im 1. Semester der 12. Klasse Oberschule mit dem KTSP überein?
3. Wie analysieren die Lehrer die Items des Deutschunterrichtstests im 1. Semester der 12. Klasse Oberschule?
4. Welche Faktoren können die Qualität der Items des Deutschunterrichtstests im 1. Semester der 12. Klasse Oberschule beeinflussen?
5. Wie ist die Qualität der Items des Tests im Deutschunterricht im 1. Semester der 12. Klasse Oberschule, Schuljahr 2008/2009?
6. Wie viele Items in dem Deutschunterrichtstest im 1. Semester der 12. Klasse Oberschule, Schuljahr 2008/2009 sind qualifiziert?

1.3 Begrenzung der Forschung

Wie ist die Qualität der Items des Deutschunterrichtstests im 1. Semester der 12. Klasse Oberschule, Schuljahr 2008/2009? Diese Forschung wird in der SMAN 99 Jakarta durchgeführt.

1.4 Problemstellung

Die Itemanalyse des Deutschunterrichtstests im 1. Semester der 12. Klasse Oberschule, Schuljahr 2008/2009.

1.5 Interesse der Forschung

Von dieser Forschung wird erwartet, dass die Vorschläge für die Deutschlehrer in der Oberschule bei der Formulierung eines besseren Deutschunterrichtstests gegeben werden.

II. ABFASSUNG DER THEORIEN UND VORSTELLUNGSGRUNDLAGE

2.1 Abfassung der Theorien

2.2.1 Begriff der Evaluation

Die Evaluation ist ein Prozess, um die Informationen zu beurteilen. Um den Erfolg des Erziehungsziels zu wissen, wird die Evaluation gebraucht. Von der Eva-

luation werden die Daten des Erziehungsziels gesammelt, um zu beweisen, es erreicht werden kann. Die Evaluation hat die wichtige Rolle im Lernprozess, deshalb soll die Evaluation gut vorbereitet werden.

2.2.2 Begriff des Tests

Der Test ist das Mittel der Evaluation. Der Test ist eine Bewertung, um das Lernergebnis und die Lernkompetenz in dem Erziehungsziel zu wissen. Ein guter Test soll objektiv, reliabel, valid, normiert, vergleichbar, ökonomisch, und nützlich sein. Es gibt einige Modelle vom Test, u.a: die offenen Aufgaben, die halboffenen Aufgaben, und die geschlossenen Aufgaben.

2.2.2.1 Begriff der Mehrfachwahlaufgaben

Die Mehrfachwahlaufgaben ist ein Test, der die mehrere vorgegebenen Antworten gibt. Dann müssen die Lernenden die richtige Antwort von den vorgegebenen Antworten auswählen. Als mögliche Antworten auf die Frage oder Fortsetzung des einleitenden Satzes werden mindestens zwei, höchstens in der Regel vier Alternativen angeboten. Nur eine der Alternative ist im allgemeinen richtig (die Antwort), alle anderen sind falsch (Distraktoren).

2.2.3 Begriff der Itemanalyse

Die Itemanalyse ist eine Analyse der Itemtests. Von der Analyse wurde es erfahren, ob er ein guter Test ist. Die Itemanalyse hat ein Ziel, den Test zu identifizieren, ob er gut, ausreichend, oder schlecht ist.

Die quantitative Itemanalyse ist eine Analyse der Mehrfachwahlaufgaben zur Identifizierung der Qualität des Items. Die Untersuchung umfasst die Berechnung des Schwierigkeitsgrads, der Trennschärfe, der Distraktoren, sowie die Untersuchung der Validität und Reliabilität.

2.2.3.1 Die Itemschwierigkeitsanalyse

Die Itemschwierigkeitsanalysen geben Aufschluss über den Schwierigkeitsgrad des gesamten Tests, dadurch kann man wissen, ob das *Item* des Tests leicht, mittelschwer, oder schwer ist.

2.2.3.2 Die Itemtrennschärfeanalyse

Die Itemtrennschärfeanalyse soll die Kriterien darüber treffen, wie gut ein *Item* zwischen guten und schlechten Kandidaten diskriminiert.

2.2.3.3 Die Distraktorenanalyse

Die Distraktorenanalyse untersucht, wie die Distraktoren angemessen funktionieren.

2.2.3.4 Die Validität

Die Validität eines Tests gibt den Grad der Genauigkeit an. Mit der Validität wird ein Test gemessen, so wird die Messung des Tests vorgegeben, und dann wird die Relevanz eines Tests zu anderem Test ausgedrückt.

2.2.3.5 Die Reliabilität

Reliabilität wird auch als Zuverlässigkeit, Genauigkeit und Stabilität der Messung bezeichnet. Sie liegt vor, wenn wiederholte Messvorgänge unter gleichen Rahmenbedingungen zu gleichen Werten führen. Ein Test kann dann als reliabel bezeichnet werden, wenn er bei einer Wiederholung des Tests unter gleichen Bedingungen und in denselben Gegenständen zum gleichen Ergebnis kommt.

2.2.3.6 Die Tabelle der Itemanalyse

Nach der Kriterien des Schwierigkeitsgrads, der Trennschärfe, der Distraktoren, der Validität, der Reliabilität können die Tabelle der Itemanalyse hergestellt werden.

2.2 Das relevante Forschungsergebnis

Das relevante Forschungsergebnis sind :

1. Eine wissenschaftliche Arbeit von Muthia Tri Handayani im Jahr 2000, nämlich Analyse der Qualität des Mathetests im 2. Quartal der 3. Klasse Mittelschule 110 Jakarta.
2. Eine wissenschaftliche Arbeit von Liya Wahyuningsih im Jahr 2006, nämlich Analyse der Angemessenheit des Indonesischetests von den Lehrern in der 7. Klasse Mittelschule nach der KBK.
3. Eine wissenschaftliche Arbeit von Kartini Megawati im Jahr 2007, nämlich Qualität des Deutschtests im 1. Semester der 10. Klasse in der SMAN 48 Jakarta, Schuljahr 2005/2006.

2.3 Vorstellungsgrundlage

Um den Erfolg des Lernergebnisses in dem Lernprozess zu erreichen, sollten die Lehrer den guten Test verwenden. Deshalb sollte die Itemanalyse von den Lehrern gebraucht werden . Durch die Itemanalyse kann es identifiziert werden, welche Items gut, ausreichend, und schlecht sind, .

In dieser Forschung geht es um den Versuch, den Schwierigkeitsgrad, die Trennschärfe, die Distraktoren,

die Validität, und die Reliabilität des Items zu analysieren. Durch die Analyse wird es deutlich, dass der Test gute Qualität hat. Ausserdem kann die Forschung auch als Evaluation von dem Lehrer benützt werden.

2.4 Hypothese der Forschung

Ha = Der Deutschunterrichtstest im 1. Semester der 12. Klasse Oberschule in Jakarta hat gute Qualität.

III METHODOLOGIE DER FORSCHUNG

3.1 Ziel der Forschung

Diese Forshung hat ein Ziel, die Qualität des Deutschunterrichtstests im 1. Semester der 12. Klasse in der SMAN 99 Jakarta Schuljahr 2008/2009 zu wissen.

3.2 Ort und Zeit der Forschung

Die Forschung wird von November bis Dezember 2008 in SMAN 99 Jakarta durchgeführt.

3.3 Methode der Forschung

Diese Forschung benutzt die kuantitative-deskriptive Methode

3.4 Population und Sampel der Forschung

Die Population und das Sample der Forschung sind 50 Items des Deutschunterrichtstests im 1. Semester der 12. Klasse in der SMAN 99 Jakarta, Schuljahr 2008/2009. Der Technik des Samples ist purposive.

3.5 Instrument der Forschung

Als Instrument wird die Tabelle der Itemanalyse benutzt.

3.6 Technik der Datenanalyse

1. Die Daten (die Antwortblätter der Schüler) werden gesammelt.
2. Die Tabellenitemanalyse wird hergestellt.
3. Die Daten werden in die Tabellenitemanalyse eingeteilt.
4. Die richtigen oder falschen Antworten jedes Schülers werden analysiert.
5. Die Daten werden mit den Formeln gezählt.
6. Das gesammelte Ergebnis der Datenanalyse wird interpretiert
7. Das Ergebnis der Analyse wird zusammengefasst.

IV. FORSCHUNGSERGEBNIS

4.1 Deskription der Daten

Die Daten der Forschung sind 50 Itemtests vom Deutschunterrichtstest im 1. Semester der 12. Klasse in der SMAN 99 Jakarta Schuljahr 2008/2009 und 350 Antwortblätter der Schüler. Die Daten werden in fünf Aspekten klassifiziert und folgt wie folgendes heraus:

Nach dem Schwierigkeitsgrads gibt es 6(12%)Items weniger als 0,30; es gibt 17(34%)Items weniger als 0,31 und mehr als 0,71; es gibt 27(54%)Items mehr als 0,71. Nach dem Trennschärfeindex gibt es 50(100%)Items weniger als 0,40. Nach der Disraktorenanalyse gibt es 25 Items (50%), die mehr als 5% sind. Dann gibt es 25 Items (50%), die weniger als 5% sind. Nach der Validität der Forschung sind 41 Items (82%) höher als r Tabelle. Und 9 Item (18%) sind niedriger als die r Tabelle. Nach der Reliabilität sind alle Wertitem (100%) mehr als 0,113 und mehr als 0,148.

4.2 Datenanalyse

Hier wird die Ergebnisse der Analyse und Klassifizierung von dem Deutschtest im ersten Semester Klasse XII SMAN 99 Jakarta Schuljahr 2008/2009 beschrieben. Die

Items werden klassifiziert und festgelegt, dass sie den richtigen Schwierigkeitsgrad, die richtige Trennschärfe, die richtige Distraktorenanalyse, die richtige Validität, und die richtige Reliabilität haben.

4.3 Interpretation der Daten

Nach der Klassifizierung des Schwierigkeitsgrads sind 12% (6 *Items*) schwierig, 34% (17 *Items*) mittelschwer, und 54% (27 *Items*) leicht. Nach der Klassifizierung der Trennschärfe sind 100% (50 *Items*) schlechte *Item*. Die Klassifizierung der Distraktorenanalyse sind 50% (25 *Items*) schlechte *Item* und 50% (25 *Items*) gute *Item*.

Nach der Klassifizierung der Validität sind 82% (41 *Items*) valid und 18% (9 *Items*) invalid. Basierend auf der Klassifizierung der Reliabilität der *Items* sind 100% der *Item* reliabel. Nach diesem Ergebnis gibt es 13 gute *Items*(26%) und 37 schlechte *Items* (74%).

4.4 Mangel der Forschung

Die Mangel dieser Forschung wird nur quantitativ analysiert. Die Qualität über die fünf Aspekten wird geprüft, nämlich: der Schwierigkeitsgrad, die Trennschärfen, die Distraktorenanalyse, der Validitäts,

und die Reliabilität. In dieser Forschung erforscht der Forscher die Fähigkeit der Schüler nicht genau.

V. SCHLUSS

5.1. Schlussfolgerung

Anhand der Ergebnisse der Analyse schließt sich heraus, dass die Validität und die Reliabilität des Deutschunterrichtstests im 1. Semester der 12. Klasse in der SMAN 99 Jakarta Schuljahr 2008/2009 gute Qualität haben, aber nach dem Schwierigkeitsgrad, der Trennschärfe, und der Disraktorenanalyse hat der Deutschunterrichtstest keine gute Qualität. Nach der Analyse gibt es die Ergebnisse, dass etwa 10(20%) Itemtests geändert werden sollen; 27(54%) Itemtests werden nicht akzeptiert; und 13(26%) Itemtests werden eine akzeptable haben oder verwendet.

5.2. Implikation

Diese Forschung hat eine Implikation für die Deutschlehrer, um die Items vom Deutschtest zu analysieren. Nach der Itemsanalyse können die Lehrer gute Itemtests wissen.

5.3. Vorschläge

Durch die Evaluierung ist zu wissen, dass es noch Mangels im gegebenen Test gibt. Die Deutschlehrer sollten die Itemsanalyse lernen, um die Qualität des gegebenen Tests zu wissen, damit sie die Mangels im gegebenen Test für nächsten Test verbessern können.

