


The Quality of Senior High School National School Examination Questions

Aloysius Mering^{1*}, Indri Astuti²

Universitas Tanjungpura, Jl. Prof. Dr. H. Hadari Nawawi, Pontianak, Kalimantan Barat 78124

¹mering7157@yahoo.co.id*; ²indribk91@yahoo.com

*corresponding author

ARTICLE INFO	ABSTRACT
<p>Article history Received January 19, 2019 Revised February 15, 2019 Accepted May 3, 2019</p>	<p>The objective of this study was to obtain information about the quality of high school National Standard School Exam (USBN) questions in Pontianak in 2018/2019 academic year. To achieve this objective, the information excavated includes (a) the process of compiling questions (b) the test of the feasibility of the questions (c) the eligibility requirements of the questions (d) administering the exam (e) the examination system of the examination scripts (f) the reporting system, and (g) the correlation of values report cards and high school USBN grades in Pontianak with UN scores. This study employed a descriptive method. The selection of high school samples used as data sources was done by random sampling and purposive sampling. Data were gathered through documentation, interviews, and questionnaires. The data were then analyzed with descriptive statistics and correlational studies. The results showed that, in the high school USBN questions in Pontianak City, (1) there were still stages in the process of drafting "not done", (2) empirical tests before the compilation of the final questions "had not been done," (3) administering the "already" applied exams as UN procedures, (4) manuscript examination system is implemented manually, (5) reporting of exam results by respective subject teachers, and (6) correlation of report card grades and high school USBN scores in Pontianak with "mostly" low "UN scores" and even in certain "negative" schools. It means that report cards and USBN scores cannot be used to predict UN results.</p>
<p>Keywords Quality Examination Analysis</p>	<p>Tujuan umum penelitian ini adalah untuk mendapatkan informasi tentang kualitas soal Ujian Sekolah Berstandar Nasional (USBN) SMA di kota Pontianak Tahun 2018/2019. Metode penelitian yang relevan dengan tujuan penelitian ini adalah metode deskriptif. Pemilihan sampel SMA yang dijadikan sumber data dilakukan dengan <i>random sampling dan purposive sampling</i>. Data diperoleh dengan teknik dokumentasi, wawancara, dan angket. Data penelitian dianalisis dengan statistik deskriptif dan kajian korelasional. Hasil penelitian menunjukkan bahwa soal USBN SMA di kota Pontianak (1) masih ada tahap-tahap dalam proses penyusunannya "tidak dilakukan", (2) uji empirik sebelum kompilasi soal akhir "belum dilakukan", (3) pengadministrasian ujian "sudah" dilaksanakan seperti prosedur UN, (4) sistem pemeriksaan naskah dilaksanakan secara manual, (5) pelaporan hasil ujian oleh guru mata pelajaran masing-masing, dan (6) korelasi nilai rapor dan nilai USBN SMA di kota Pontianak dengan nilai UN "sebagian besar" "rendah" dan bahkan pada sekolah tertentu "negatif". Artinya nilai rapor dan USBN tidak dapat dipakai untuk memprediksi hasil UN.</p> <p>This is an open access article under the CC-BY license.</p> 

I. Introduction

Schools determine high school students' graduation in Indonesia through teacher board meetings (BSNP Regulation Number: 0048/BSNP/XI/2018) with graduation criteria

(1) completing all learning programs, (2) obtaining a minimum "good" attitude/behavior value, (3) taking the National Examination, and (4) passing the National Standard School Examination (USBN). One of the graduation

criteria is passing the National Standard School Exams. National Standardized Exams run the standards set by the National Education Standards Agency (BSNP), such as the percentage (20%-25%) of questions from the center (national) included in USBN questions through the application. The rest, 75%-80% of USBN questions, are arranged by the subject teacher consultation forum (MGMP). Fulfillment of school exam standards so that they are considered to meet the national standard criteria (USBN) is another meaning of the achievement of the quality requirements of USBN questions according to the standards set by the national examination manager, the Education Assessment Center (PUSPENDIK). (Hattie, J & Brown, G, 2010) say, assessment refers to 'What's so?' And evaluation to 'so what?' Both depend on high-quality measurement, and both focus on the qualities, degrees, and characteristics of student learning of the material deemed necessary by society and identified in the curriculum. Fulfillment of these quality standards is reflected in (1) how the USBN question is organized, (2) how it is administered, and (3) how the USBN results are examined and processed. According to (Brown, H, 2000), there are three qualities of a test; they are practicality, reliability, and validity. These three test quality requirements can be fulfilled through the procedure of preparing systematic tests and empirical tests. If following general procedures, the preparation of learning outcomes tests such as USBN includes steps (a) developing test specifications, (b) writing questions, (c) reviewing questions, (d) assembling questions (for testing purposes), (e) test trials, (e) item analysis, (f) item selection and assembly (final form), (g) test printing, (h) final form test administration, and (i) scale and norm arrangement (Suryabrata, 2000). Likewise, administering the exam requires standardized preparation and supervision so that the development of exam questions that have been appropriately pursued is not in vain. The last essential part is the examination of examination texts and reporting. This section demands very high integrity from teachers and schools. It was concluded that the quality of the question scripts, administration, examination/reporting of USBN was essential because the results of USBN were one of the instruments for determining student graduation. (Flateby, 1996) explained, some of the possible purposes of classroom achievement test and their re-consultation are to (1) measure an individual's achievement of course ob-

jectives, (2) assess the group's performance, (3) evaluate the test and the items, (4) evaluate and improve instruction and the curriculum. The results of the research are carried out through a systematic process. As explained in the (Capsim management simulations, 2018), it is explained that learning outcome assessment refers to the systematic process of collecting and analyzing data about student learning and using this information to improve degree programs and educational quality continuously. (Capsim management simulations, 2018), the organizer of the exam must understand that each step of the examination management determines the quality of the questions constructed. Thus, all those involved in the process and organization of the test should give high appreciation to the test set that was made, because the test set will ultimately place a person, group (institution), process (business), and curriculum or program in a specific quality category. Since USBN was reinstated in 2017, there must be research examining the implementation of the USBN. Information from several high school teachers said that the analysis of USBN questions was carried out just after the exams were conducted. Therefore, the validity and level of confidence in the test results need to be tested.

Information about the quality of the Pontianak City High School USBN questions in this study was obtained through internal examination of questions to get data about the level of difficulty of the items, different items, the functioning of the distractors, and the editorial in accordance with the standard test rules. In addition, the relationship between report card grades and USBN was analyzed with the National Examination (National Examination) scores as criterion scores.

II. Method

In accordance with the objectives of the study, namely, to obtain information about the quality of high school USBN questions in Pontianak in 2018/2019 Academic Year, the research method used was descriptive research Fig. 1.

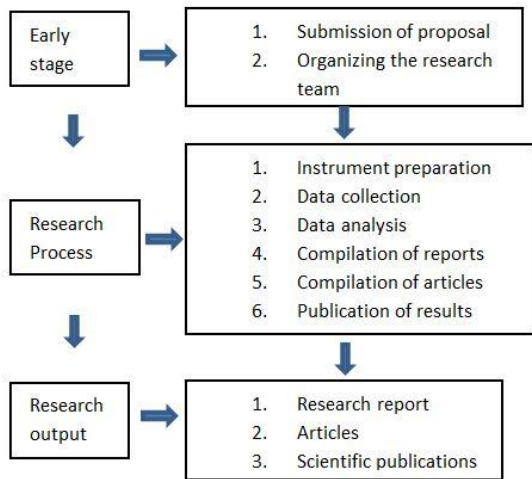


Fig. 1. Research Process

A. Research sample

This research was conducted in Pontianak City High School in 2018/2019. All schools can be sampled. All schools that meet the main requirements, namely analyzing the USBN manuscript independently after the test, are sampled. The selection of sample high schools also considers school qualifications. However, many schools do not save the original value of the USBN results. Schools that meet the requirements are chosen randomly (random sampling) and appointment (purposive sampling). The selected schools are SMA X and Y (accreditation qualifications A) and SMA Z (accreditation B).

B. Research Stage

The research process will take place from 20 May to 20 November 2019 through the following stages. The research process chart is as follows:

1) Observed aspects, data collection techniques, and data analysis

With regard to the USBN data of Pontianak City High School in 2018/2019, the aspects of being observed were (a) the process of compiling the questions, (b) the procedure for the feasibility of the questions, (c) the questions about the feasibility of the questions, (d) administering the exam, (e) examination script examination system, (f) USBN results reporting system, and (g) correlation of report card grades and high school USBN scores in Pontianak with UN scores. Data from research aspects were obtained by documentation techniques, and grading scales filled out by the teacher. The research data were analyzed with descriptive statistics and correlational analysis.

III. Result and Discussion

Research on USBN Pontianak city 2018/2019 academic year directed (1) the process of compiling the questions, (2) the test of the feasibility of the questions, (3) the eligibility requirements of the questions, (4) administration of USBN, (5) the USBN manuscript inspection system, (6) the results reporting system, and (7) the correlation of report card grades and USBN scores with UN scores. Information Number 1 through Number 6 is obtained through a value scale instrument that is answered by respondent teachers, while problem data Number 7 is from school documents.

Preparation, due diligence, quality requirements, implementation, the examination of manuscripts, and reporting system on the results of USBN from the respondent teacher's answer about "the process of compiling the USBN questions in Pontianak City High School" it is known that (1) the lattice preparation of USBN questions is "always" made, (2) the preparation of questions is done by the teacher team, (3) item analysis is carried out after USBN, (4) the empirical test of the USBN question is "not done", (5) the assembly of the USBN question is carried out by the drafting team based on a logical study. But from the results of the study of the items that the researchers did use the test analysis program by (To & Wibisono, 2003), it turned out that "the feasibility test for high school USBN questions in Pontianak" was conducted after the test (only for reporting and administration test) or "not done before the test" as a condition for standardizing the test. Based on the analysis of item (a) there is no test format and language of the questions, (b) the distribution of abnormal levels of difficulty, (c) the power of different items is not analyzed, because about 40%-50% items do not meet the requirements of good power difference, (d) about 95% of items have distraction that do not work (category "bad" to "very bad"), and (e) the average USBN item has test items that do not meet the validity requirements, except Mathematics. The results of the analysis of Pontianak City USBN items in 2018/2019 school year are summarized in Table 1.

Table 1. Summary of Results of Differential Power Analysis, Difficulty Levels, Format Analysis, Significance, and Deception Functions

Question Item Analysis		Indonesian		English		Mathematics	
		Science	Social	Science	Social	Science	Social
Discrimination Power	-Good	22	32	32	33	28	23
	-Not Good	18	8	8	7	2	7
Difficult	Easy	28	26	22	16	6	3
	Medium	8	11	12	14	18	19
	Difficult	3	3	5	10	6	8
Significance	Significant	4	13	22	23	21	20
	Not Significant	36	27	18	17	9	10
Format and Language analysis	Done	Yes	Yes	Yes	Yes	Yes	Yes
	Was Not Done	-	-	-	-	-	-
Distraction	Work	0	4	5	5	6	6
	Doesn't work	40	36	35	35	24	24
Conclusion		revision	revision	revision	revision	revision	revision

^a Indonesian and English = 40 Multiple Choice questions

^b Mathematics = 30 Multiple Choice questions

Reviewing the eligibility requirements through reviewing USBN question items in Pontianak city high school in 2018/2019, according to the respondent's teacher has been conducted on the "format and language of the questions," "the level of difficulty of the questions," "analysis of the power of different items," "key functioning and deception," and "item validity" was rated as "very good." However, the results of the analysis of the items by SMA Y showed that the item analysis was carried out after the implementation of the USBN, and the results showed that all the USBN question sets should be "revised" and almost all questioners (distractors) "did not work." For the implementation of USBN, it was considered "very good" by the respondent teacher.

The system of checking the USBN script according to the respondent teacher is done manually by the teacher of each subject.

Likewise, the documentation of the pure value of the USBN results is also done by the subject teacher. In this study, researchers only had access to analyze the pure value of USBN results in SMA Y.

Reports on USBN results are not only known by subject teachers, but also by students, schools, and all related parties, including the pure grades of USBN. A list of grades (form/blank form) USBN has also been provided uniformly for all schools.

Correlation of Report Card Score (X1) and High School USBN (X2) value in Pontianak with UN (Y). The correlation format between the Report Card Value (X1) and the USBN (X2) value of SMA in Pontianak and the UN (Y) is summarized in Table 2.

A summary of the results of correlation analysis between the Report Card Value (X1) and the USBN (X2) value of SMA in Pontianak with a UN (Y) value can be seen in Table 3.

Table 2. Correlation Formats (R_{x1y} and R_{x2y}): Report Card' Value-UN; USBN-UN value

School	Major	Subject					
		Indonesian		English		Mathematics	
SMA X	Science	R_{x1y}	R_{x2y}	R_{x1y}	R_{x2y}	R_{x1y}	R_{x2y}
	Social	R_{x1y}	R_{x2y}	R_{x1y}	R_{x2y}	R_{x1y}	R_{x2y}
SMA Y	Science	R_{x1y}	R_{x2y}	R_{x1y}	R_{x2y}	R_{x1y}	R_{x2y}
	Social	R_{x1y}	R_{x2y}	R_{x1y}	R_{x2y}	R_{x1y}	R_{x2y}
SMA Z	Science	R_{x1y}	R_{x2y}	R_{x1y}	R_{x2y}	R_{x1y}	R_{x2y}
	Social	R_{x1y}	R_{x2y}	R_{x1y}	R_{x2y}	R_{x1y}	R_{x2y}

Table 3. Correlations (R_{x1y} and R_{x2y}): Report Cards-UN; USBN-UN value

School	Major	Subject					
		Indonesian		English		Mathematics	
		Report & UN (R_{x1y})	Report & UN (R_{x2y})	Report & UN (R_{x1y})	Report & UN (R_{x2y})	Report & UN (R_{x1y})	Report & UN (R_{x2y})
SMA X	Science	0,5266	0,4445	0,276	0,046	0,266	0,128
	Social	0,57	0,284	0,141	0,005	0,266	0,067
SMA Y	Science	0,325	0,4348	0,475	0,740	0,738	0,718
	Social	0,5736	0,45707	0,5875	0,5803	0,4404	0,3903
SMA Z	Science	0,22	0,321	0,416	0,465	0,23	0,189
	Social	-0,108	-0,118	-0,592	-0,34	0,316	0,384

Based on the categorization of correlation (Trihendradi, 2009), 0.7–1.00 (high); 0.4–0.70 (sufficient or substantial); 0.2–0.40 (low); <0.2 (very low or neglected), it can be concluded the relationship: Value of Report-UN (r_{x1y}); The value of USBN-UN (r_{x2y}) of Indonesian, English, and Mathematics subjects in the majors of Natural Sciences and Social Sciences (SMA X, SMA Y, and SMA Z) in Pontianak in the 2018/2019 academic

year showed "negative", "low" correlation levels and "very low". In other words, report cards and USBN scores cannot be explained or used to predict UN scores. If you look at the pattern of these relationships, the report card and USBN values tend to be "very homogeneous" (around 8), while the UN value is more heterogeneous between 40-100 can be seen in Fig. 2.

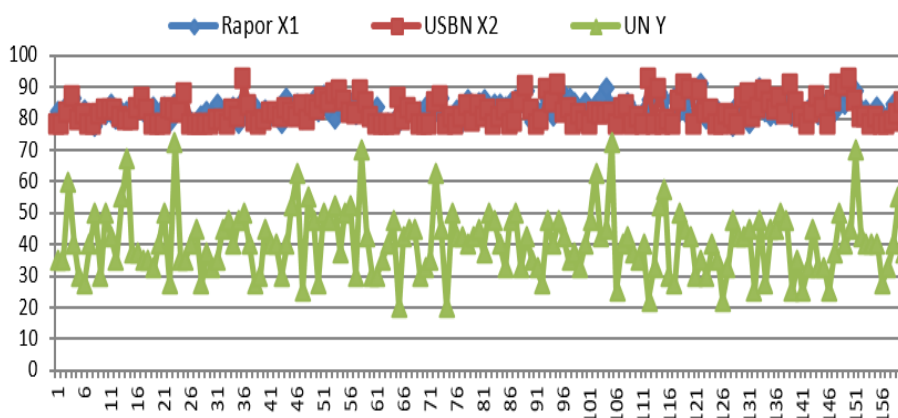


Fig. 2. Example of Relationship of Report Card Value, USBN, and National Examination in Mathematics Subject of Senior High School X Department of Natural Sciences

In the Y science majoring school Y, the correlation report card grades-UN; the value of USBN-UN Indonesian subjects is "very low," meaning it cannot be explained or used to predict UN scores, while the correlation score of UN report cards is categorized as "sufficient". For English Subjects, the value of USBN is highly positively correlated with

UN scores. Likewise, Math scores, both report cards, and grades are positively correlated with UN scores. If you look at the pattern of relationships between the report cards-USBN-UN and the three relationships are very close and positive. The range of values and value movements are more "homogeneous" (between values 60-100). Look at Fig. 3.

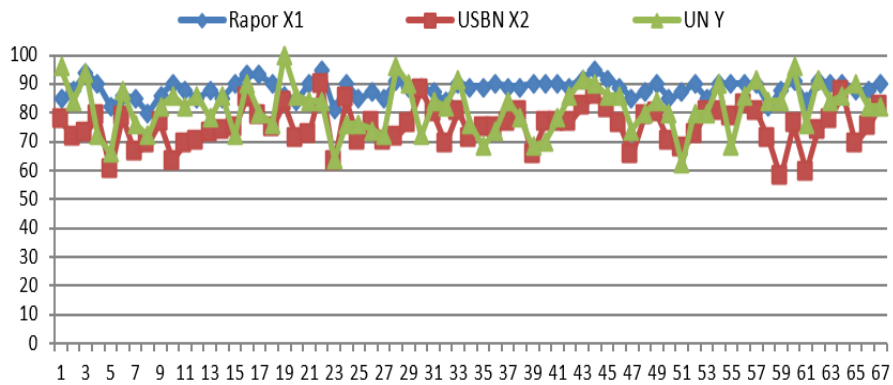


Fig. 3. Example of Relationship of Report Card Value, USBN, and National Examination in Indonesian Language Subjects in SMA Y Department of Natural Sciences

For SMA Z, the correlation report-UN and USBN-UN Indonesian subjects have a "low" correlation value. For English subjects, the correlation scores are "sufficient," and in Mathematics, the correlation scores are "low" and "very low." It causes the report card and USBN values "cannot" be used to base UN predictions. That is, the high-low value of

report cards and USBN, there is no relationship and contribution to the value of the UN. The tendency of report cards and USBN scores is homogeneous (around 8) and separated from UN scores (Indonesian between 30s-80; English between the 20s-60s; and Mathematics between 20s-30s). Look at Fig. 4.

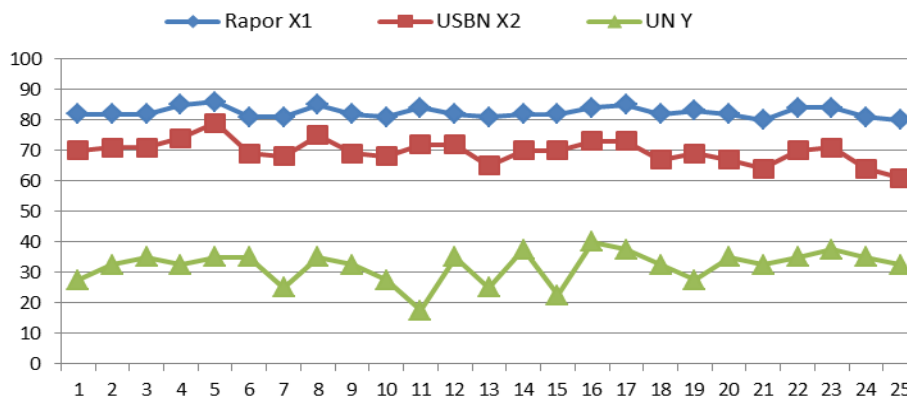


Fig. 4. Example of Correlation of Report Card Value, USBN, and National Examination in Mathematics Subject of SMA Z Department of Natural Sciences.

Correlation level (R_{x1y} and R_{x2y}): Report card-UN score; The USBN-UN scores of Indonesian, English, and Mathematics subjects in the Social Sciences majors (SMA X, Y High School, and Z High School) of Pontianak City in 2018/2019 Academic Year can be explained that (1) the correlation of Report Card Value with Indonesian Language UN categorized as "sufficient" ($r = 0.570$), (2) the correlation of USBN scores with UN in Indonesian subjects was categorized as "low" ($r = 0.284$), (3) the correlation of Report Score with English national examination subjects was categorized as "very low" ($r = 0.141$), (4) the correlation of USBN scores with UN in English subjects is categorized as "very low" ($r = 0.005$), (5) the correlation of report card grades with UN in Mathematics subjects

is categorized as "low" ($r = 0.266$), (6) the correlation of USBN scores with UN Mathematics subjects is categorized as "very low" ($r = 0.067$).

All correlation values ("very low"; "low"; and "enough") in SMA X cannot explain the contribution of report cards and USBN scores to the UN. The unexplained thing is the source of report cards and USBN (whether the value is pure or not). When viewed from the correlation graph, the report card and USBN values are in a range of homogeneous values (between 80s-90s) and separate from the UN value. It means that the report card grades and USBN are almost unrelated to UN scores, especially in Mathematics (there is no correlation). Look at Fig. 5.

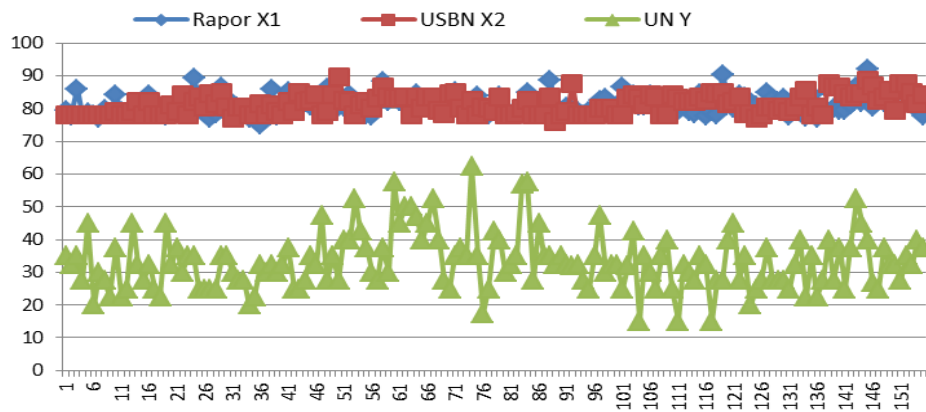


Fig. 5. Example of Relationship of Report Card Value, USBN, and National Examination in Mathematics Subject of Senior High School X Social Sciences Department

For SMA Y, the correlation between the Report Card Score and USBN with the UN in Indonesian, English, and Mathematics subjects majoring in Social Sciences is in the "sufficient" category, except the correlation of USBN scores with the UN for "low" Mathematics subjects. Each correlation value (1) The report card grades with the Indonesian national examination are categorized as "sufficient" ($r = 0.574$), (2) The USBN scores with the Indonesian national examination are categorized as "sufficient" ($r = 0.571$), (3) Report cards with English national examination are categorized as "sufficient" ($r = 0.588$), (4) USBN scores with English national examination are categorized as "sufficient" ($r = 0.580$), (5) Report grades with Mathematics national examination are categorized "Enough" ($r = 0.440$), and (6) USBN scores

with UN Mathematics subjects are categorized as "low" ($r = 0.390$).

All correlation values ("enough" and "low") in SMA Y have not been able to significantly explain the contribution of report cards and USBN scores to the UN. When viewed from the correlation graph, the report card values are more likely to be higher than USBN and UN values. The relationship between the value of USBN and UN is very close so that the value of USBN can be used to predict UN results. Specifically, in Mathematics subjects, the report cards are separated from the USBN and UN scores. All values (report card, USBN, and UN) show the same motion (linear). Look at Fig. 6.

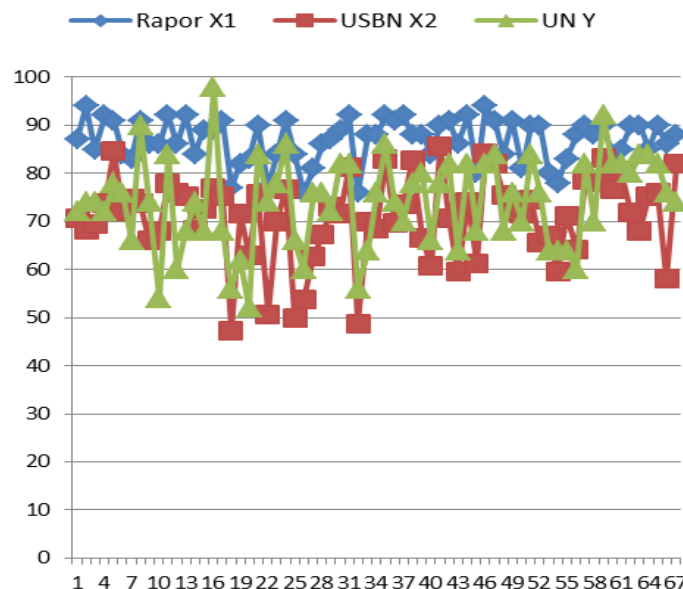


Fig. 6. Example of Relationship of Report Card Value, USBN, and National Examination in Indonesian Language Subjects in SMA Y Social Sciences Department

For SMA Z, the correlation between report card scores and USBN with the UN in Indonesian, English, and Mathematics subjects majoring in social studies is in the "negative" category, except the correlation of USBN scores with the UN for math subjects is "low". Each correlation value (1) The report card grades with the Indonesian national examination subject are categorized as "negative" ($r = -0.108$), (2) The USBN scores with the Indonesian national examination subject are categorized as "negative" ($r = -0.118$), (3) The report card grades with the English language UN are categorized as "negative" ($r = -0.592$), (4) The USBN scores with the English language UN are categorized as "negative" ($r = -0.340$), (5) The report card grades with the

UN Mathematics subjects are categorized as "low" ($r = 0.316$), and (6) USBN scores with UN Mathematics subjects are categorized as "low" ($r = 0.384$).

All correlation values ("negative" and "low") in SMA Z cannot explain the contribution of report cards and USBN scores to the UN. Correlation value "negative" means "every increase in student's grade" on the report card grades and USBN followed by a "decrease in each student's national examination." When viewed from the correlation graph, the report card, USBN, and UN values tend to be separate, uncorrelated, or do not contribute to each other. All three values (report card, USBN, and UN) show the same motion (linear). Look at Fig. 7.

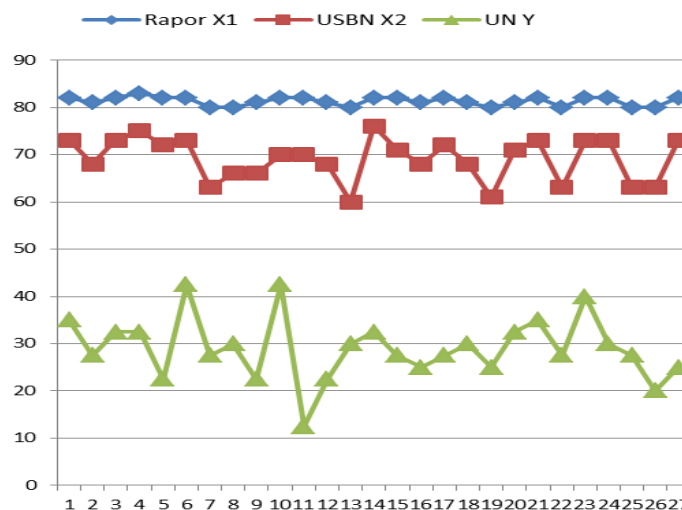


Fig. 7. Example of Relationship of Report Card Value, USBN, and National Examination in Mathematics Subject of SMA Z Social Sciences Department

If you examine the comparison of the average score of report cards, USBN, and UN (Fig. 8). in Indonesian, English, and Mathematics (for example the Department of So

cial Sciences), the most "inconsistent" value relation is SMA Z, SMA X "less consistent," and SMA Y tends to be "consistent."

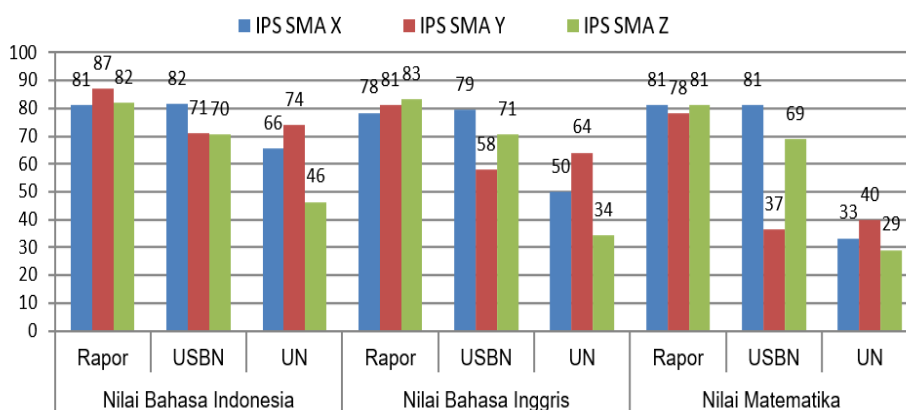


Fig. 8. Comparison of Average Report-USBN-UN Score in SMA X, Y School, and Z School of Social Sciences

Preparation, due diligence, quality requirements, implementation, the examination of manuscripts, and reporting system on the results of USBN. The process of compiling the USBN questions that have been carried out by the USBN question compilers includes (a) the compilation of the questions grid, (b) the making of questions by the drafting team, (c) there is an internal trial run, (d) there is an analysis of the questions, and (e) peer review. This process has partially met the requirements for the preparation of standard questions (Suryabrata, 2000) However, due diligence in the sense of empirical test after the preparation of the questions "has not been done," so that the question user does not have valid information about the quality of USBN questions. Overall, the feasibility test for USBN questions is "often" done. For USBN questions which are included in general tests, they must meet the eligibility element to be able to "provide accurate and accurate information" (Azwar, 2016). Information from the respondent's teacher was confirmed by the results of the item test using the Test Analysis program by (To & Wibisono, 2003), which showed that "the feasibility test for high school USBN questions in Pontianak" was carried out after the test. That is, based on the analysis of items (a) the distribution of the level of difficulty is not normal, (b) the power of different items is not analyzed, because about 40%-50% items do not meet the requirements of good power difference, (c) about 95% items have a deception that does not work (the category of "bad" to "very bad"), (d) the average USBN question set has test items that do not meet the validity requirements, except Mathematics.

From the assessment scale instruments, all teacher respondents stated, "the format and language of the questions," "the level of difficulty of the questions," "analysis of the different items," "key functioning and deception," and "item validity" - "very good." The opinions of these respondent teachers were "incompatible" with the results of the analysis of the items conducted. From the analysis of items, it was concluded that all USBN test items should be "revised," especially deceiving questions that "don't work much." Thus, the matter of high school USBN in Pontianak in the 2018/2019 school year "did not meet the standards".

The implementation of USBN Pontianak City High School in 2018/2019 "concerns (a) instructions for working on USBN, (b) com-

pliance and supervisory behavior, (c) compliance with USBN schedule, and (5) supporting facilities. The information provided by the respondent's teacher stated that "the implementation of USBN Pontianak City High School in 2018/2019 academic year" was "very good." Supervision of the implementation of the exam plays an important role, because good questions, carefully constructed will be useless if the oversight factor is ignored. Proper supervision of exams avoids participants from "dishonest" actions, as contended by (Sudijono, 1998), that "evaluation of learning outcomes can be expressed as a good evaluation if it can be separated from factors that are subjective."

The examination system for USBN question scripts is done (a) by the subject teacher himself, (2) scoring objective questions (multiple choice) by the subject teacher himself, (3) the scoring and essay question weighting are determined by the subject teacher himself, (4) the examination is done in a manner manual, and (5) the setting and use of USBN pure grades are determined by the teacher according to school direction. This problem creates a gap between USBN grades between schools. Schools that aim to use USBN scores as material for self-quality study and evaluation will tend to prioritize pure values, while schools that solely use USBN grades as learning report data tend to process USBN grades to "satisfy" users. It is what causes "why low or medium qualified school grades — the USBN scores tend to be higher than schools with" good or very good "qualifications (opinion of LJ, schoolteacher Y). This fact is contrary to the findings of (Pratiwi et al., 2016) that learning outcomes reflecting the competency that he gained. With this fact, the essay question cannot be analyzed by the researcher because the weight varies or is determined by each teacher.

Reporting on the results of the USBN is carried out by each subject's teacher. This report is submitted to schools, students, parents, and the Provincial Education Office. In this case, the honesty of the school is required to make a report. The report, if objective, will be very useful for feedback for improving overall education programs (Nitko, 1996). Predictably, the report card grades and USBNs of the two high schools that were respondents to the study were processed products that "deviated" from the true values (see graph in the findings section). This illustrates as in (Elmunyah, 2014) findings that national

education policies will work well if teachers and school administrators can work together to achieve the success of curriculum goals.

A. Correlation of Report Card Value (X1) and USBN value (X2) with UN value (Y)

The correlation between report card grades and USBN values with UN scores is important to explore. In this case, the UN value is the criterion value that becomes the standard, especially for USBN scores. The higher the correlation between USBN scores and UN scores, the better the USBN questions compiled by subject teachers. If the USBN value is not positively correlated, then the USBN seems to be implemented to meet the demands of formality (meeting regulations). Especially if there are teachers or schools that "don't appreciate" the value of USBN by changing the results (control), then USBN does not mean anything to students, parents, schools, and subject teachers. USBN, as a public exam (for high school students), should meet the quality standards of good questions, which have passed the different power tests, difficulty level, validity, and the test of the deception function.

In the case of USBN Pontianak City High School in 2018/2019 for Indonesian, English, and Mathematics, the item analysis was carried out after the implementation of USBN to fulfill the completeness of the report. An empirical test for standardization of questions that should be carried out before the test. Items of empirical test results are compiled into USBN question sets. From the results of the post USBN analysis, it turns out that (a) there is no format and language test for USBN questions, (b) the distribution of abnormal levels is not normal, (c) the different items are not analyzed, because about 40% - 50% items do not meet the requirements good difference, (d) about 95% of items have distraction df that do not work ("bad" and "Very bad" categories), and (e) the average USBN test kit has test items that do not meet the validity requirements (not significant), except Mathematics problems. The failure to carry out the feasibility test process on USBN means that it is not in line with the opinion of the assessment experts, as stated by (Hartiningtyas et al., 2016; Setyabudi, 2011; Sudijono, 1998), that a good learning achievement test at least fulfills valid, reliable, objective, and practical requirements so that it can be used to make decisions about students, as (Popham, W, 2010) also dis-

closed (Bhandari & Lingzay, 2014) that appropriate educational decisions depend on the accuracy of educational assessments because, quite obviously, accurate assessments will improve the quality of decisions whereas inaccurate assessments will do the opposite.

IV. Conclusion

Teacher learning outcomes tests, school exams, USBN, and UN, are all tests conducted by teachers (teacher-made tests), schools (semester exams, school exams), and government (USBN and UN). At any level, the test instrument must meet the standards of good tests conducted through due diligence, such as (a) the format and language test, (b) test the difference in grain power, (c) test the level of difficulty of the item, (d) the validity test of the item, (e) reliability testing, (f) equivalent scoring, (g) administration test, (h) processing of results, and (i) reporting system. From the analysis conducted, it was found that (1) the process of compiling USBN high school questions in Pontianak in 2018/2019 academic year was not standardized or not "fully" following the systematic compilation of questions, (2) empirical tests or feasibility tests before the final question compilation had not been done, (3) the administration of the exam has been carried out as a UN procedure, (4) the script inspection system is carried out manually by each subject teacher, (5) the weighting of essay questions turns out to be determined by each subject teacher, (6) the USBN results reporting system implemented by the subject teachers respectively to students, teachers, schools, parents, and local government institutions, (7) the correlation of the report card grades and the USBN high school scores in Pontianak with the "most" UN scores reflect the "weak" correlation, "very weak" and "negative", except for SMA Y ("sufficient" correlation), and (8) "Generally" UN scores are much lower than report cards and USBN values.

From these findings, it was suggested to the USBN Manager (the Office of Education and Culture of West Kalimantan Province, the Principal Work Meeting, and Subject Teachers' Consultation) that USBN be prepared through a standardized question construction process. The process of compiling USBN questions should go through the stages of test planning, question compilation, due diligence, and compilation of questions after the due diligence. To achieve the conditions of objec-

tivity and honesty (fairness), the system of supervision and inspection must also be equal and standard.

References

- Azwar, S. (2016). *Tes prestasi*. Pustaka pelajar.
- Bhandari, R., & Lingzay, T. (2014). Vocational maturity of senior secondary school students in relation to their family environment. *Scholarly Research Journal for Interdisciplinary Studies*, 2(14), 2138–2145.
- Brown, H. D. (2000). *Principles of language learning and teaching* (Vol. 4). Longman New York.
- Capsim management simulations. (2018). *The ultimate guide to learning outcomes assessment*.
- Elmunsyah, H. (2014). A national education policy-based ICT model for Indonesian vocational high schools (VHS). *Global Journal of Engineering Education*, 16(3), 136–140.
- Flateby, T. L. (1996). A guide for writing and improving achievement tests. *Tampa: University of South Florida, Office of Evaluation and Testing*.
- Hartiningtyas, L., Purnomo, P., Elmunsyah, H., & Nurmalasari, R. (2016). Correlation between creative thinking skill and competency achievement with vocational maturity on vocational high school. *AIP Conference Proceedings*, 1778(1), 30032.
- Hattie, J. A., & Brown, G. T. (2010). Assessment and evaluation. In *Educational psychology: Concepts, research and challenges* (pp. 116–131). Routledge.
- Nitko, A. J. (1996). *Educational assessment of students*. Eric.
- Popham, W. J. (2010). *Everything school leaders need to know about assessment*. Corwin Press.
- Pratiwi, A. S., Sudjimat, D. A., & Elmunsyah, H. (2016). Contribution of industrial work practice performance and creativeness to the academic skill and its effect to the outcome of skill competency test of computer and network technology skill package in vocational high school. *Journal of Education and Vocational Research*, 7(4 (V)), 31–36.
- Setyabudi, I. (2011). Hubungan antara adwersiti dan inteligensi dengan kreativitas. *Jurnal Psikologi Esa Unggul*, 9(01), 126306.
- Sudijono, A. (1998). *Pengantar evaluasi pendidikan*. PT RajaGrafindo.
- Suryabrata, S. (2000). *Development of psychological measurement tools*. Andi.
- To, K., & Wibisono, Y. (2003). Petunjuk instalasi dan pengoperasian Anates versi 4. *Jurusan Psikologi Pendidikan Dan Bimbingan Universitas Pendidikan Indonesia*.
- Trihendradi, C. (2009). Melakukan analisis statistik menggunakan SPSS 17. *Yogyakarta: C. V Andi Offset*.