



Public Opinion on National Exam Policies in Indonesia

Danny Manongga^a, Ade Iriani^b, Sutarto Wijono^{c,*}

^aFakultas Teknologi Informasi, Universitas Kristen Satya Wacana

^bFakultas Teknologi Informasi, Universitas Kristen Satya Wacana

^cFakultas Psikologi, Universitas Kristen Satya Wacana

Keywords :

Educational Policy,
National Examination,
Online News,
Public Opinion,
Sentiment Analysis

Abstract : Every new policy by Indonesian government in National Examination (NE) implementation always obtains different respond from public. Since the implementation, NE system already experienced many changes, but in recent years this system receives serious critiques. As a result, government then abolished this system as graduation determinant in 2014. This research analyzes public opinion, in the form of positive and negative sentiment toward NE policy, and factors that drive the opinions. Data in this research obtained from online news media from 2012 to 2015. The result shows that public sentiment fluctuating from year to year and depends on three important factors, i.e. political pressure, extreme events, and media coverage.

1. Introduction

Indonesian Ministry of Education and Culture (MEC) continued under the spotlight in the society because of National Examination (NE) policies. NE became important function to determine student graduation in elementary and junior high school. NE has been applied for a long time. Since the beginning, this system had been through several changes in its name and policies without drawing serious critique from the public. But since 2012, this system received many critiques from the public. The critique reached the peak that the government had to cancel NE as graduation determinant in 2014.

Public opinion has important impact to the direction of government policy. Baum (2004) explained how public scrutiny can reduce the use of force in the case of Operation Restore Hope. While, Brooks & Manza (2006) conclude that public opinion have strong influence to the generosity of social programs in rich democratic countries.

However, public opinion really depends on information supply from news media. In other words, news media, especially online news media, plays key role in forming the public opinion because besides providing news, this media also receiving public opinion directly. In addition, this media provides

large data to facilitate researchers to investigate public opinion towards various government policies. Several researchers using online data to public opinion research, i.e. (Anstead and O'Loughlin, 2014; Bright et al., 2014; Burger, 2011; Dickinson and Hu, 2015; Sha et al., 2014; Das et al, 2014; Khan et al., 2014; Vilares et al, 2015). However, regardless of the availability of the large data, public opinion analysis towards government policies in education field still limited caused by two main reasons. First, opinion analysis emphasizes on detecting expressions, emotions, viewpoints and private states, expressed in contents (Pang & Lee, 2008; Wiebe et al., 2004). Second, the task in news domain is more complicated than to those in reviews. Many newspapers want to give an impression of objectivity so that journalists will often refrain from using clearly positive or negative words, such as expressions are subtler, and news quotation is often shorter (Balahur et al., 2009). In addition, unlike opinion analysis in English that well investigated, opinion analyses in non-English language like Indonesian have not much to be investigated. Sukhum (2011) shows the difficulty in opinion analysis in Thai language.

Current research aims to analyze public

* Corresponding authors

e-mail addresses : danny.manongga@staff.uksw.edu, ade.iriiani@staff.uksw.edu, sutartown@yahoo.com

opinion toward NE policies based on articles in online news media. Two main questions become research focus: 1) how public opinion about high school NE; 2) what factors drive public opinion towards NE?

The rest of this article consists of five sections. The first section provides background and evolution of standardized testing in Indonesia. The second section provides a literature review outlining standardized testing and sentiment analysis. The third section describes research methodology and data. The fourth section presents empirical findings of opinion analysis. The final section provides the discussions of findings.

2. Development of National Examination in Indonesia

NE in Indonesia have unexpected policy changes. Schools and students often become the victim of those changes, like the determination of test subjects, graduation standard and even the impact of student who failed the NE. In general, NE format that had the longest period was EBTANAS (National Final Stage of Study Evaluation) for 22 years, while the shortest was UAN (National Final Exam) for just two years. These changes did not indicate innovation, but rather showed the problem in the existing educational concept (Suyitno, 2013). Further explanation will review the NE's format changes since 1965.

In 1965 to 1971, final test used standardized test called Ujian Negara (State Exam) and applied for all subjects. The central government run the exam and applied uniformly to all regions in Indonesia. Also, the government strictly supervised the test, so that graduation percentage only reached 50%.

In 1972 to 1979, government gave freedom for every school or school group. The government only developed general guidelines, while school implemented the final test with loose monitoring that graduation percentage reached 100%.

In 1980 to 2002, government decided two forms of final test, EBTA (Final Stage of Study Evaluation) and EBTANAS. Central government performed EBTANAS for general subjects, while province government performed EBTA for non-EBTANAS subjects. Schools had role in determine graduation through a combination of the two forms of the exams plus the value of daily tests that were listed in the report card. Students could graduate if the average grade of all tested subjects is six (see in Table 1), even though there were one or more subjects had grade under three (Nurfuadah, 2014).

Table 1 NE graduation grade changes since 1980 to 2015

Year	Minimum grade	Minimum average
1980-2002	-	6.0
2003	3.01	6.0
2004	4.01	-
2005	4.25	4.25
2006		4.50
2007	5.00	5.00
2008	4.25	5.25
2009		5.50
2010		
2011	4.00	
2012		
2013		
2014		
2015	Graduation determined by school	

Source: summarized from various source

Ideally, EBTANAS aimed to control and improve the quality of education and to obtain a uniform indicators for comparisons between schools. However, EBTANAS implementation have many weakness, such as:

- 1) It could not measure student's academic achievement comprehensively;
- 2) It just tested student's ability temporally in short time;
- 3) It just collected information about student's cognitive ability; and
- 4) It reduced the learning process to become just exam question practices.

In 2003-2004, because of weaknesses in EBTANAS, government then replaced it with UAN. The main difference between EBTANAS and UAN is the way to determine student's graduation. In EBTANAS, graduation was determined by combination of grades in first semester, second semester and EBTANAS. While in UAN, graduation was determined by individual subject grades. UAN had different graduation standard in each year. In UAN 2003 (see Table 1), graduation standard was 3.01 for each subjects and minimum average grade was 6.00. Test question made by central government and school could not lift UAN grade. Student who failed still had chance to repeat the test in a week after UAN. In UAN 2004, government decided graduation standard to 4.01 for each subject. The government also removed the minimum average grade and the re-exam for students who failed. However, shocked by poor UN result, government then created a conversion table to increase student's grade. This table turned out introduced a big unfair element. Student who answer correctly for more than half questions obtained lowered in their final score to pull other student's grades below (Syahril & Lesko,

2007). Furthermore, society also gave pressure to the government to held re-exam to students who failed in first test.

In 2005 to 2010, government replaced UAN to UN and increased graduation standard in each year (see Table 1). In UN 2005, government set the minimum grade for each subject to 4.25. Students who failed in first test could take second test for failed subjects. In UN 2006, government set the minimum graduation standard to 4.25 for each tested subject and the average should be above than 4.50 and there was no re-exam. In UN 2007, government set the minimum grade for each subject to 4.25 and the average minimum 5.00 for all the subjects. If there a minimum grade 4.00 in one of the tested subjects, the other two subjects must reach grade 6.00. Government also removed re-exam for failed student, but held Package C Exam (National High School Equivalency Examination) or repeated the UN test next year. In UN 2008, government increased the number of tested subjects from three to six subjects. It is done because students tend to be more concerned with subjects tested in UN than other subjects. Government still maintained minimum grade from 2007, but increased the minimum average to 5.25. In UN 2009 and 2010, government increased the minimum average to 5.50, but allowed two subjects with grade 4.00.

In 2011 to 2014, government set the graduation based on combination of school grades and UN grades with percentage of 60% for UN grades and 40% for school grades. Graduation grades for each subject minimum 4.00 and minimum average 5.50. In 2012, government came up with idea that UN as a replacement for SNMPTN (National Entrance Test for State Universities). With the UN as part of SNMPTN, then there were some questions that equated to a questions in SNMPTN. Government argued that this policy would make SNMPTN be more practical, economic, and free of irregularities. For that reason, government involved universities in arrangement of UN questions and demanded UN should be honest and credible (Kompas, 2012). However, that idea faced challenge from several universities leader (AntaraJatim.com, 2012).

In order to made UN to be honest and credible, government then planned to increase the number of question variation from five to ten types in a class. In 2013, government applied UN with twenty types of question. This created chaos in UN implementation because the complexity of printing and logistic distribution to all over Indonesia. As a result, eleven provinces had to cancel the execution UN in their region. This cancellation inflicted serious critique from public to the government.

In 2014, after the selected president had appointed, the new minister of education ends the UN polemic with remove UN as graduation requirement.

In 2015, UN had no longer to be graduation determinant for student, but just for mapping. Grades in the diploma was the combination of 60% report grades (first to fifth semester) and 40% school test made by local Department of Education.

3. Literature Review

3.1. Review of public opinion towards Standardized Testing

The key question about UN is what factors encourage the emerge of public opinion toward government policy. Various literature around this problem grouped in three main factors: 1) political pressure, 2) media coverage, 3) extreme events.

3.2. Political Pressure

Political pressure could make school treat students in a discriminatory manner. Keddie (2012) stated that the application of standardized test as accountability and audit tool for both national and international has generated a sense of urgency in schools to increase student education achievement. School then categorize their students based on their test result. Students who are in the higher category will enjoy various privileges. In the other hand, students who are in the lowest category considered unprofitable or become deficit for school (Ball, 2010, p155), therefore it is not attractive for school (Muijs et al, 2011, p86). Gillborn and Youdell (2000) show “educational triage” practices where school try to maximize “return” of their limited sources. Ball, Maguire and Braun (2012) give deep insight about how “policy enactment” from performative policy gives pressure which characterizes this economic form, along with testing and valuation practice in school. The same thing also said by Hardy (2014) that policy enactment as a product of differential relationship existed between they who try to dominate field of schooling practices and how this thing play it roles in practice. Alike with students, political pressure impacted to teacher’s profession threatened to be commercial. Teachers must be able to add “value”, or increase institution reputation where they were working (Keddie, 2012). School administrator judge the teachers based on a set of excellent indicator and the comparison with the target. Demand of performative environment embed a terror feeling to the teachers (Ball, 2003), while distrust state and continuous monitoring through agent and other institution threatened teacher’s autonomy and professionalism (Lingard & Sellar, 2012).

Those stressful conditions created paradox inside the teacher (Keddie, 2012; Ball, 2003). For some teachers, that condition raise self’s autonomy and professionalism, but for others, that condition develop inner conflict, inauthenticity and resistance - a game process. Teachers learn to play the game in a cynical fashion. As expression of a teacher in Ball (2003, p220):

We become ontologically insecure: unsure whether we are doing enough, doing the right thing, doing as much as others, or as well as others, constantly looking to improve, to be better, to be excellent. And yet it is not always very clear what is expected... We are unsure what aspects of work are valued and how to prioritize efforts. We become uncertain about the reasons for actions. Are we doing this because it is important, because we believe in it, because it is worthwhile? Or is it being done ultimately because it will be measured or compared? It will make us look good!

Some researchers had shown a lot of undesirable consequences proof that impacted in students learning (Thompson & Harbaugh, 2013; Thompson, 2013). The same consequences happened in US, Great Britain, and Australia which experienced in the implementation of standardized test. Negative practices that happened like teaching only for the upcoming test and narrowing curriculum focus (Klenowski & Wyatt-Smith, 2012), or create less inclusive class environment (Comber, 2012; Comber & Nixon, 2009; Lingard, 2010; Polesel et al, 2012; Thompson & Harbaugh, 2013).

3.3. Media Coverage and Public access to accurate scientific information

Media coverage gives clear public understanding towards the impact of standardized test policy. For example, publication of NAPLAN experience in Australia for last five years makes public understand the reality that the best student's achievement result included in the regular rate (ACARA, 2012b). Statistically, improvement did occur in Year 3 Reading, Year 5 Reading and Year 5 Numeracy, but there was no significant national improvement in other categories (ACARA, 2012a; Thompson, 2014).

Besides, that publication had make the public gave different opinion. For example, publication of NAPLAN test result in MySchool site since 2010 generated diverse opinion from various academic community. For some party, this information received positively because it enabled people to do comparison among schools based on their test result. However, for the others, especially the teachers, this publication actually gave negative impact to them because it reduced their complex job became a target numbers (objective) or student's achievement standard. Teachers then became a passive consumer of all information produced by things outside their daily life (Hardy & Boyle, 2011).

3.4. Bad News that affect Public Opinions on A Standardized Testing

Bad news could affect public opinion change to standardized tests in four essential things. First, news about valuation failure. In 2008 England's key stage 3, national curriculum tests for seven and fourteen-years-old have been erased because of

failure in valuation caused delay in test result reception for thousands of students (Mansell et al, 2009). The same thing also happened with Key Stage 2 science SATs (Standard Assessment Tests) in May 2009. This occurred following a boycott of the tests by many head teachers. Teachers feel concerned with the impact of those tests to children education (Welcome Trust, 2011). Teacher assessment based on coursework, practical work, and fieldwork has long been an integral part of GCSE courses, but because of concerns about malpractice, the teacher-assessed components are restricted and eventually removed in most subjects (Ofqual, 2014). Second, test implementation which deviated from its initial purpose. The initial purpose of standardized test implementation is to measure standard achievement in order to improve public school's quality and to inform parents about educational progress of their children. But in its development, the number of test should be held by school has increased significantly that many school hours spent on the test preparation and implementation (Nelson, 2013). Besides, usage of test result that far deviated from its original purpose, as to assign school grades and passing score (Strauss, 2012), cause the post publication of test result become the most turbulent times for public schools and for the education profession (Spar, 2012). Third, the implementation cost that increases from year to year. The number of test and implementation cost could be bad news not only for school and student, but also for government. For example, the annual budget allocated by the Texas State Legislature for the 2013 fiscal year was over \$86 million (State of Texas, 2013). These costs represent only those at the state level for the creation, distribution, scoring, and reporting of the results. However, the majority of the costs of implementing tests falls on the local schools (Phelps, 2000; Crow, 2014). In Florida State, usage of computer-based testing protocols also requires the purchase of additional computers, as well as increased bandwidth, both of which must come out of the district operating dollars. For schools, the practice tests must be also printed and copied by each individual school. While for students who did not reach standard score in FCAT class 10 will lose their chance to choose because they must attend the remedial class. All costs, preparation, testing and reaction to test result, accumulated to whole system cost (Strauss, 2012). Finally, test result become bad news for teacher and school future. Test result often followed by monetary consequences to teacher and school income. For teacher, test result could lead to work dismissal (Strauss, 2012). Bradley & Fryer (2011) give details of incentive impacts for both teacher and school in several school districts.

3.5. Review of opinion analysis

In computer science, opinion analysis has various term, such as opinion mining or sentiment

analysis. Those names reflect the focus of opinion analysis in detecting the polarity and emotion recognition (Liu, 2010; Liu, 2012; Cambria et al, 2013). In the beginning of the development, this branch of science focuses in customer sentiment to products and movies. In that analysis, the text under analysis only cover one product or one film so that sentiment can be easily classified as the positive or negative (O'Keefe et al, 2013; Balahur & Steinberger, 2009).

Since the beginning of its development, opinion analysis has experienced many progress, including the development of its methods and tools (Pang & Lee, 2008; Cambria et al., 2013). Nonetheless, there are several reasons that this area still requires further research. First, studies of opinion analysis still focused on subjective types of text such as blogs, twitter, or reviewing a product or film. The authors of the text types are mostly express their opinion quite clear. News articles have received less attention, even though they have many sources. News articles and media reports typically contain less clear expressed opinions. Although there is support or criticism, but bias or sentiment of journalists often expressed indirectly, for example by highlighting some facts and ignoring others (Balahur & Steinberger, 2009).

Some early attempts that concentrate on news articles, among others were (Fortuna et al., 2009; Belyaeva and Van der Goot, 2009). However, research that analyzes public opinion toward government policies in education based on online news is still limited. News articles about government policies become a useful target for opinion mining, as they discuss the salient opinion by people whom newsworthy. In addition, the news articles often provide accurate quotations attributed to the opinion of the speaker. Nevertheless, there is a challenge to a quote from a topic of debate, since the polarity of the target and the meaning is less clear (O'Keefe et al., 2012).

4. Methodology

4.1. Data Collecting Technique

Kompas.com media become the source for UN news. This site chosen because it has credibility in delivering the objective news. News about UN collected from 2012 to 2015, but limited from January to June. This limitation has been done because UN held in April and the result announced in May, so news content after June will no longer contain relevant information. However, in 2012 to 2015 had been chosen because public gave various opinion about UN in these periods.

The search and news selection utilize search engine provided by Kompas.com. The search process began with finding news points. In every news usually contain several news points which have link with the others. So, when searching process finds a news point, the next searching will just follow

the other link to catch news provided in the link. This process repeated for several times until all links have been explored and there no more news.

4.2. Data Processing Technique

News article saved in separated file based on month and year. This process has purpose to find UN issues or themes that became headline in particular month. While to analyze the opinion or behavior, all news from January to June saved in same file.

4.3. Government Policy Analysis

Public opinion is a response to government policies. This research identifies many government policies about UN which triggered public opinion. Government policy could have formal form like constitution and unwritten policy. Unwritten policy could have form like government declaration or government behavior which sometime conflicted with formal policy. In this research, government policy is every news in media which contain government declaration, such as: education and culture official, like Menteri Pendidikan dan Kebudayaan (Mendikbud), Badan Penelitian dan Pengembangan (Balitbang), Badan Standar Nasional Pendidikan (BSNP), or local education department.

4.4. Public Opinion Analysis

Analysis to public opinion done by classify various public statements, complaints, sentiments, or behaviors toward UN in positive and negative category. Public opinion that did not contain one of those categories will considered as neutral or just give fact without opinion.

Public opinion toward UN have wide dimension. Classifying process to two categories above faces many challenges, because news article often contains direct quotation, indirect quotation, and sometime does not contain any clear sentiment. To minimize error, this research adopted annotation guide from Balahur and Steinberger (2009) which conclude as below:

- Only focus on sentiment expression about certain part, not all of the text.
- Identify news item which clearly contains positive or negative opinion.
- Annotation without involve specific knowledge about the speaker. If there is uncertainty, let it be neutral or objective.

Example of positive, negative and neutral annotation:

Positive: "Selection to PTN will be more practical with using UN grades as one of the determinant."

Negative: "Do not force the UN. The one who should be fixed is not the UN, but the quality of education, teacher and student."

Objective: "Credible UN result become one of the requirements to enter Perguruan Tinggi Negeri

(PTN). Sell power to PTN not about UN quality, but the quality of education process.”

● Do not annotate bad news and good news. For example: “UN implementation in eleven provinces delayed caused by logistic trouble.” This statement only contains fact, even if there is negative news content.

5. Result

5.1. Opinion Holders

Public opinion cannot be separated from opinion holder. Opinion holder is an individual or group who give their statement on news media. In fact, opinion about UN came from wide circle of society. Because of it, opinion holder then grouped

into categories which have same background. For example, Education and Culture Minister and Vice Minister, BNSP, Balitbang, and Education and Culture General Inspectorate came from same institution, so they could be grouped in one category. While non-governmental organization(LSM) represents various opinion from Teacher Forum, Indonesian Corruption Watch, Advocacy Team, Education Coalition, Education Practitioners, courses and other mass organization. This grouping produce nine different categories of opinion holder. However, this grouping did not vanish opinion difference from each individual in certain group. Table 2 shows the nine categories.

Table 2 Nine categories of opinion holder

No	Opinion Holder	Participants
1	Presidency	President and Vice President
2	Ministry of Education and Culture (Kemendikbud)	Minister and Deputy Minister of Education and Culture, National Education Standard Agency (BSNP), Research and Development Agency, and General Inspectorate
3	Local Government	Governor, Regent/Mayor, Chief Department of Education, Regional Education Board
4	College	Rector of University, Vice Rector, Lecturer
5	Legislative	Central and Regional Legislative Assembly
6	School	Headmaster, Teachers, Examination Committee
7	Non-governmental organization (LSM)	Teachers Forum, Indonesian Corruption Watch, Advocacy Team, Education Coalition, Education Practitioners, Courses, Companies, Mass Organization (PBNU)
8	Parents	
9	Students	

5.2. Aggregate Trend of Public Opinion

After mapping opinion holder into nine categories, the next step is input annotation positive or negative opinion from individual into the right category. This opinion annotation done per year to find public opinion development from 2012 to 2015.

Table 3 shows the public opinion annotation result from 2012 to 2016 based on nine categories before. Generally, this table shows that government dominates online news with their opinion.

Table 3 Development of Public Opinion from 2012 to 2015

No	Opinion Holder	2012			2013			2014			2015		
		JA	P	N	JA	P	N	JA	P	N	JA	P	N
1	Presidency	-	-	-	-	-	-	-	-	-	4	4	-
2	Kemendikbud	29	29	-	38	26	5	18	18	-	27	27	-
3	Local Government	17	17	-	24	12	14	10	10	-	15	15	-
4	College	8	3	6	10	3	7	3	3	-	2	2	-
5	Legislative	8	-	8	6	2	4	1	1	-	-	-	-
6	School	9	8	1	18	3	15	9	9	1	13	13	-
7	LSM	8	-	8	32	1	31	1	-	1	2	-	2
8	Parents	3	3	-	-	-	-	-	-	-	-	-	-
9	Students	6	6	-	8	-	8	5	1	4	6	6	-
	Total	88	66	23	136	47	84	47	42	6	69	67	2

Abbreviation: JA= sum of article; P= sum of positive opinion; N= sum of negative opinion. While objective opinion was not included.

Figure 1 shows the development of public opinion from 2012 to 2015. In 2012 there were 88 articles (JA) in online news which contain 66 positive sentiments and 23 negative sentiments. The number of article still limited under 2013, but far across the number of article in 2014 and 2015. It caused by in 2013, government introduces UN issues as the replacement of SNMPTN which received response from various circle in society, especially from college, legislative and LSM.

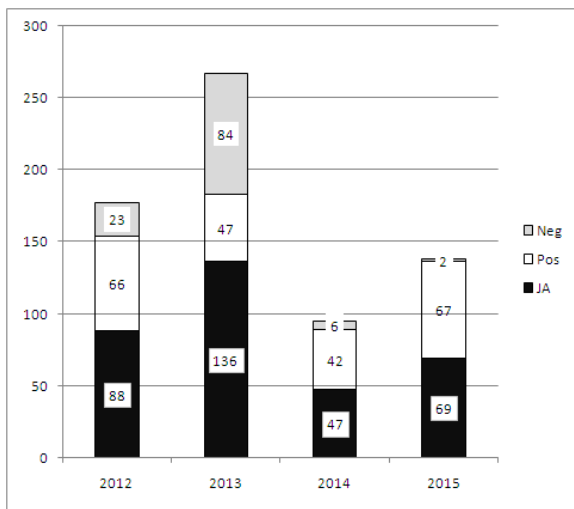


Figure 1 Opinion aggregate from 2012 to 2015

In 2013, the number of article in online media increases from 88 to 136 articles, while the number or negative sentiment increase from 23 to 84, and the positive sentiment decrease from 88 to 47. The logistic problem and delay of UN execution in eleven provinces from thirty-three provinces in Indonesia became the main issue of the raise of negative sentiment and the decrease of positive sentiment. In this year government increase the number of question from five types to twenty types of question. This addition cause difficulty in question printing process and lateness in logistic delivery in all over Indonesia.

In 2014 there was no extreme event. Learned from previous year, government had prepared enough time to print the question and deliver the logistic. In the same year, general election was held so the society attention focused more on the president candidates than issues about UN. In 2015 the new Minister of Education under the lead of President Joko Widodo, erased UN as the graduation determinant, but also introduce computer-based UN. This received response from various circle, but principally there was no negative sentiment toward UN implementation.

Figure 2 shows opinion aggregate trend with narrowing the nine categories to be only two categories, which are governmental and non-governmental. Category 1 to 3 in Table 3

included in governmental, while category 4 to 9 included in non-governmental.

Generally, the sum of article (JA) contain government opinion are greater than non-governmental opinion, except in 2013. The difference impacted also in the number of positive opinion which always came from governmental group than the non-governmental. In the other side, the number of negative opinion usually came from non-governmental group. The government always try to dominate public opinion through statements in online news media that support the importance of UN. But government effort did not successful in 2013 when non-governmental group gave a lot of critiques of chaos happened in UN implementation. In 2013 when the chaos occurred, the number of negative opinion (N) not only greater than the positive (P) one, but in fact half of negative opinion came from the government itself, like Education and Culture Ministry and Local Government (see Table 3). Negative opinion which came from Education and Culture Ministry mostly contain “tossing responsibility” statements.

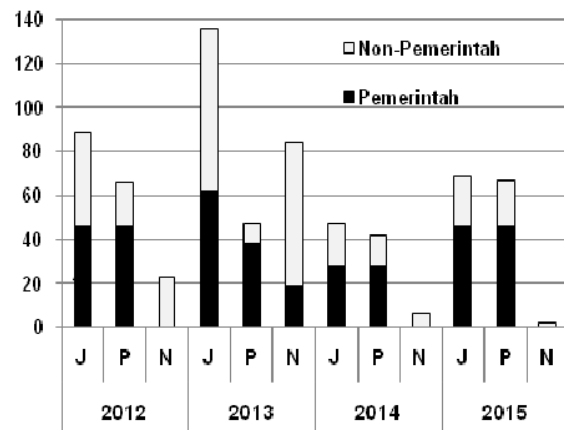


Figure 2 Opinion aggregate based on governmental and non-governmental

5.3. Attributes that Push Public Opinion toward UN

Public opinion toward UN did not far from attributes and factors that triggered those opinions. Same like opinion holder identification, public opinion toward UN have diverse attributes, so it need to be classified based on several attribute categories. Table 4 shows five UN attributes which draw most of public attention.

Based on these five attributes, UN question got the most attention from public. This attribute has almost flat number of opinion from 2012 until 2015 with 930 opinions or 37.18% of total opinion. Graduation takes second place with 607 opinions or 24.27% of total opinion. Cheating and demand of honest UN take third place with 570 opinions or

22.80%, while two other attributes which have less attention from public are computer-based test and UN delay.

Table 4 Attribute which draw many public opinions

No	UN Attributes	2012	2013	2014	2015	Total	%
1	UN question	233	380	144	173	930	37.18
2	UN graduation	193	227	107	168	695	27.79
3	Cheating and demand of honest UN	224	101	49	196	570	22.8
4	Computer-based UN	12	7	17	169	205	8.2
5	UN delay	2	96	2	1	101	4.03
	Total	664	811	319	538	2501	100

Specifically, public opinion toward UN question consist of three important things, those are: difficulty level, the confidentiality, and the diversity of question. Question difficulty level divided by three: 1) hard, medium, and easy; 2) the implementation of same level of question for all over Indonesia; and 3) the increase of question difficulty every year. The question confidentiality including about centralized question printing in Jakarta or decentralized to each province. Centralized printing has obstacle like question distribution and security that make the local region did not have good responsibility. While if decentralized printing, there is concern that every region has the chance to increase their local graduation level. The diversity of question including the number of question variety in a class. In 2012, the number of public opinion toward UN question reached 233 opinions because government made five different types for 20 students in a class. In 2013, government decided to increase the five types into twenty types of question. This addition caused the increase of UN funds, which already high and in the same time increased the difficulty in printing and distribution process to all over Indonesia. The consequence is the lateness logistic delivery that delayed UN in eleven provinces.

Public opinion toward UN graduation consist of three things: 1) UN as the graduation determinant for students, 2) student and school preparation, and 3) graduation impact. UN as the student graduation determinant get many opinions because it becomes the source of other opinions. Discussion around this topic related with the authority equality of central government in form of UN weight and school authority to determine student graduation. This debate also connect with UN function, is it for education quality mapping or UN as student final evaluation. Second topic related to student graduation is student and school preparation to success in UN graduation. Most school plan their preparation for students feel ready to face the UN. Since student enter the third grade, learning dynamic also change because it will determine the success level when facing the UN. So, the school must pour

their maximum energy to sharpen student's ability to answer the UN question which typically in multiple choice form. School also give addition study hours outside regular study hours for UN subjects and hold try out several times. Besides, school also prepare students mentality, such as give relaxation and motivate the students to become their best at national examination (UN). Third topic related to graduation is the impact of graduation to the students and related party. Student graduation have wide impact, not only for the students and parents, but also for the teacher, headmaster, Head of Ministry of Education, even the Mayor and Governor. For students, the pressure feels really intense that appear in various expression to celebrate their graduation like signing their uniform with marker and convoy their vehicle across the road. In the other side, graduation level has great impact for teachers, headmaster, and Head of Local Ministry of Education career. If the graduation level is low, then the teacher or headmaster threatened to be transferred to less popular school. While the for Head of Local Ministry of Education, low graduation level could lead to lost his position. Even the Head of Region could be considered fail to grow higher education level in his region that could lead to loss of society trust to be elected again.

UN graduation attribute get the most public opinion in 2013 because of two reasons (see Table 4). First reason is in this year there were delayed of UN execution in eleven provinces. This delay not only worsen the stress level of students and teachers in UN preparation, but also leads to concern that the number of students who graduated will decreased. Second reason is the delay shows that UN management only project oriented than education quality enhancement. This lead opinion that the increase of UN question from ten types to twenty types only inflated UN budget, but did not upgrade education quality and not reduce the cheating level happened in UN implementation. In addition, this policy is contradictive with government statements that always deny the cheating acts in UN.

Third attribute with the most opinion is cheating in UN and demand of honest UN implementation. The total opinion for this attribute is

570 and the highest opinion occurred in 2012. In this year, the government introduce an idea of vertical integration in all stages of education, starts from primary school until college. Vertical integration aims to use certain stage test result as the consideration to student admission in higher stage. This idea means UN result could be used to replace SNMPTN. This idea received many public opinions. College considered UN and college admission selection have two different function. Other than that, they feel that UN have many weaknesses, especially about cheating and grade mark up by schools. They point to the reality that students grade from suburban schools have better result than popular schools.

To make college receive UN result, the government try to reduce cheating in every level which draw public attention. Government efforts such as asking the student and UN controller to signed a declaration to perform UN honestly, recruited independent test controller from lecturer in colleges, increase the number of question type in a class, asked college to store test logistic and mobilized policeman to guard and distribute the question for schools. Besides the government, school also have ways to press cheating level, such as forbid the students to bring cell phone inside the test room, check students with metal detector, and install CCTV to monitor test room. But those efforts receive various reaction by public because it did not reduce the cheating action in UN implementation. All level in the society asked those efforts because only symptomatic and did not solve the main problem. Cheating happened because government implement the same standard without considering condition of each region in Indonesia. This implementation triggered the students, teacher, and headmaster to act cheat. In the other side, government attitude shows different facts, in one side they disagree with cheating act, but in the other side their action indicates cheating practices.

Fourth attribute which draw public opinion is computer based UN. In 2015 the government introduces computer based UN. Government aim to minimize the cheating act in national examination, also press the cost for implementation of UN that increases from year to years. This effort received positive response, but still have not implement together yet because there were many schools that did not have enough electricity, computer, and internet facility.

The last attribute is delay of UN which happened in 2013 (see Table 4). This opinion raised because in 2013 the government increased question types from five to twenty types. This addition caused problem in lateness of logistic delivery that impacted to UN delay in eleven provinces. This delay impacted to society demand the Education Ministry to be responsible for that matter and proposal to audit the government tenders about the way they provide the UN material.

Table 4 shows in 2014, the number of public opinion toward UN drastically sinking. It because of the general election that held in the same year, even the date also near with UN implementation. Besides, campaign content of one strong candidate indicates the review of UN policy, even the proposal to erase UN as the school graduation determinant.

6. Discussion

Public opinion toward UN came from wide circle in the society. However, the government try to dominate the news coverage in various online media, but it does not block people to give their opinion as the balancer opinion. This happened as the great impact of UN.

The number of article in 2012 to 2015 show fluctuation trend based on issues or policy came from the government and received many responses from society (see Figure 1). Generally, the number of positive opinion mostly dominate the negative opinion, except in 2013. In this year, the number of UN article reached highest record as the impact of extreme events. In addition, only in this year the number of opinion from non-governmental group overshadowed the article contain government opinion (see Figure 2). Figure 2 shows that positive opinion mostly came from government level, while negative opinion usually came from non-government people.

Generally, there are three main factors drive public opinion toward government policy related to UN, such as political pressure, extreme events and media coverage. Political pressure to students, teachers, headmasters, and the Head of Local Education Ministry came in various form. First related to the high proportion of UN in determining student graduation (60% of UN and 40% of school test) which cause stress level for the students and trigger cheating act. Second, UN brings up moral hazard in the circle of educator and students because it has dominant roles to determine graduation. UN considered as a key to determine the progress and access for student to take higher education level. As the consequences, cheating appeared more systematically that supported by teacher, school, and local government who want higher graduation percentage. Third, the failure has great impact to all party prestige, so all students, teachers, and school energy poured with drilling method that will sharpen student's ability to answer UN questions which generally in the form of multiple choices. Because of that drilling method, analysis and literature ability being sacrificed. Fourth, due to UN graduation grade will determine the acceptance in college, so school marked up the school test grade. In the end, same UN standard for all regions in Indonesia bring new injustice between regions and schools. Disadvantage regions and schools in infrastructure and human resource should accept the reality as the risky victim to leave the "game arena".

UN result usage practices pushed the occurrence of cheating by teacher, headmaster, and Head of Local Education Ministry. First, the decline of school graduation level will decrease school prestige in front of the society and will impact to the decreasing of student number, also the fund assisting from government to school. For state schools, decline of graduation level will impact to teacher and headmaster transfer to less popular school. For Head of Local Education Ministry, it could threaten his position if the graduation level is low. Government political pressure creates chain effect not only for students who attend the test, but widens to all parties outside the school. Even the central government also feels anxious if the UN graduation level is going down. Parliament (DPR) and society will evaluate that government generally and especially Ministry of Education who could not increase national education quality. So, on one side, there is government will to increase education standard quality, but on the other hand there is anxiety if the UN graduation level decreases from the previous year.

In above situation, non-governmental public opinion has a portion in driving government policy change. Public opinion especially when occurred together with general election, gives enough political pressure to be heard by the new government that UN role as graduation determinant finally be erased.

UN delay shows extreme event, where eleven from thirty-three provinces in Indonesia cannot be held on the fixed date. Besides, the new governmental issues or policies regarding UN implementation give impact to the increase of public opinion. For example, the addition of question types from five to twenty causing increase of UN cost. This cost increase became public and Parliament (DPR) attention in the middle of public antipathy toward UN domination to increase education quality, even could not reduce the cheating act.

Other extreme event is the government eagerness to integrate all education stages through UN. UN has been used as a requirement in admission of higher level education. For example, the higher-level education using UN result to filter students who register in the school. However, it only happened in primary school until high school. For college admission, the selection process using admission screening test. After several observations, the government willing to erase admission test to state college and replace it with UN result. But it receives many critiques from many levels of college. They believe that UN and admission test are two different things, because UN considered as learning evaluation in certain level of education, while admission test is an evaluation to aim the student to choose the right major in college. Every major has its own criteria to determine the acceptance of students, so it cannot consider equal with UN grade. Besides, college hesitates the UN grade can reflect someone's real ability because UN grades is the combination of

UN grades and school test grades. They also think why suburban schools can reach better grades than other well-known schools in society.

Pay close attention to challenges in college, the government then include college in every UN implementation process, such as question arrangement, storing and distributing logistics, as well participate in UN monitoring. With that participation, the government try to widen UN issues which begin with primary to high school problems to be higher level issues. This way, the government actually try to blur the UN purpose itself, to change from student final learning evaluation to be college admission test.

Media coverage gives chance to the public to get information directly and accurate about UN implementation. Media availability makes public be able to follow every development of government policies. online news media also give chance to the public to express their opinion directly and free, so it could be alternative opinion.

7. References

- [1] Balahur, A., B., Steinberger, R., v. d. Goot, E., Pouliquen, B. & Kabadjov. M. (2009). Opinion Mining on Newspaper Quotations, Proceeding of Int. Joint Conference on Web Intelligence and Intelligent Agent Technology, 3, 523-526.
- [2] Anstead, N., & O'Loughin, B. (2014). Social media analysis and public opinion: the 2010 UK General Election. *Journal of Computer Mediated Communication*, online . pp. 1-17. ISSN 1083 -6101 .
- [3] AntaraJatim.com. (2012, January 17). PTN Tolak UN jadi Syarat SNMPTN (State Universities reject national examination as replacement for SNMPTN). Retrieved June 20, 2015, from AntaraJatim.com: <http://www.antarajatim.com/lihat/berita/80604/ptn-tolak-un-ja-di-syarat-snmptn>
- [4] Balahur, A., & Steinberger, R. (2009). Rethinking Sentiment Analysis in the News: from Theory to Practice and back. Troyano, Cruz, Díaz (Eds.). WOMSA'09, (pp. 1-12).
- [5] Ball, S. M. (2012). *How schools do policy: Policy enactments in secondary schools*. London: Routledge.
- [6] Ball, S. (2010). New Class Inequalities in Education: Why Education Policy maybe Looking in the Wrong Place! *Education Policy, Civil Society and Social Class*. *International Journal of Sociology and Social Policy*, 30 (3-4), 155-166.
- [7] Baum, M. A. (2004). How Public Opinion Constrains the Use of Force: The Case of Operation Restore Hope. *Presidential Studies Quarterly*, 34 (2).
- [8] Belyaeva, E., & van Der Goot, E. (2009). News bias of online headlines across languages. The study of conflict between Russia and Georgia. *Rhetorics of the Media*. Conference Proceedings. Lodz University Publishing House.
- [9] Bradley, M., & Roland, G. F. (2011). *The Power and Pitfalls of Education Incentives*. Washington: The Hamilton Project.
- [10] Bright, J. M. (2014). The Use of Social Media for Research and Analysis: A Feasibility Study. Retrieved November 3, 2015, from UK Government-Department for Work: <https://www.gov.uk/government/organisations/department-for-work-pensions/about/research#research-publications>
- [11] Brooks, C., & Manza, J. (2006). Social Policy Responsiveness in Developed Democracies. *American Sociological Review*, 71, 474-494.
- [12] Burger, J. D. (2011). Discriminating Gender on Twitter. Proceedings of the Conference on Empirical Methods in Natural Language Processing, (pp. 1301-1309).
- [13] Cambria, E., Schuller, B., Xia, Y., & Havasi, C. (2013). New avenues in opinion mining and sentiment analysis. *IEEE*

- Intell. Syst, 28 (2), 15-21.
- [14]Comber, B. (2012). Mandated literacy assessment and the reorganisation of teachers' work: federal policy, local effects. *Critical Studies in Education*, 53 (2), 119-136.
- [15]Comber, B., & Nixon, H. (2009). Teachers' work and pedagogy in an era of accountability. *Discourse:Studies in the Cultural Politics of Education*, 30 (3), 333-345.
- [16]Crow, J. E. (2014). An Analysis of the District Marginal Costs of Mandatory State Standardized Tests" (2014). University of Nebraska-Lincoln, Educational Administration: Theses, Dissertations, and Student Research. Paper 231.
- [17]Das, T. K., Acharjya, D. P., & Patra, M. R. (2014). Opinion Mining about a Product by Analyzing Public Tweets in Twitter. *International Conference on Computer Communication and Informatics (ICCCI -2014)*. Coimbatore, INDIA.
- [18]Dickinson, B., & Hu, W. (2015). Sentiment Analysis of Investor Opinions on Twitter. *Social Networking*, 4, 62-71 Published Online July 2015 in SciRes.<http://www.scirp.org/journal/sn>.
- [19]Fortuna, B., Galleguillos, C., & Cristianini, N. (2009). *Detecting the bias in media with statistical learning methods Text Mining: Theory and Applications*. Taylor and Francis Publisher.
- [20]Gillborn, D., & Youdell, D. (2000). *Rationing education: Policy, practice, reform and equity*. Buckingham: Open University Press.
- [21]Hardy, I. (2014). A logic of appropriation: Enacting national testing (NAPLAN) in Australia. *Journal of Education Policy*, 29 1: 1-18., 29 (1), 1-18.
- [22]Hardy, I., & Boyle, C. (2011). My School? Critiquing the abstraction and quantification of Education. *Asia-Pacific Journal of Teacher Education*, 39 (3), 211-222.
- [23]Keddie, A. (2012). Matters of autonomy and accountability in the English schooling policy context: constraints and possibilities. CPPR lunchtime seminars at King's College, London. Retrieved April 20, 2015, from <http://www.kcl.ac.uk/sspp/departments/education/research/cp pr/Seminars/seminars.aspx>
- [24]Khan, F. B. (2014). TOM: Twitter opinion mining framework using hybrid classification scheme. *Decision Support Systems*, 57, 245-257.
- [25]Khan, F., Bashir, S., & Qamar, U. (2014). TOM: Twitter opinion mining framework using hybrid classification scheme. *Decision Support Systems*, 57, 245-257.
- [26]Klenowski, V., & Wyatt-Smith, C. (2012). The impact of high stakes testing: The Australian story. *Assessment in Education: Principles, Policy and Practice*, 19 (1), 65-79.
- [27]Kobayashi, N. e. (2004). Collecting evaluative expressions for opinion extraction. *Proceedings of theInternational Joint Conference on Natural Language Processing (IJCNLP)*.
- [28]Kompas.com. (2012, January 9). Sekolah Siasati Dana UN (Schools use tactics to fund the national examination). Retrieved July 23, 2015, from Kompas.com: <http://properti.kompas.com/read/2012/01/19/03070174/sekola h.siasati.dana.un>
- [29]Lingard, B. (2010). Policy borrowing, policy learning: testing times in Australian Schooling. *Critical Studies in Education*, 51 (2), 129-145.
- [30]Lingard, B., & Sellar, S. (2013). Catalyst data': Perverse systemic effects of audit and accountability in Australian schooling. *Journal of Education Policy*, 28 (5), 634-656.
- [31]Liu, B. (2012). *Sentiment Analysis and Opinion Mining*. Morgan and Claypoo.
- [32]Liu, B. (2010). Sentiment analysis and subjectivity. In *Handbook of Natural Language Processing*, Second Edition. CRC Press, Taylor and Francis Group.
- [33]Liu, B. (2010). Sentiment Analysis: A Multi-Faceted Problem. *IEEE Intelligent Systems* , 1-5.
- [34]Mansell, W., James, M., Group, A. R., & Paul. (2009). *Assessment in schools. Fit for purpose? A Commentary by the Teaching and Learning Research Programme*. London: Economic and Social Research Council, Teaching and Learning Research Programme.
- [35]Mujis, D., Aincow, M., Chapman, C., & West, M. (2011). *Collaboration and Networking in Education*. London: Springer.
- [36]Nelson, H. (2013). *Testing More, Teaching Less.What America's Obsession with Student Testing Costs in Money and Lost Instructional Time*. Washington: American Federation of Teachers.
- [37]Nurfuadah, R. N. (2014, December 31). Ujian Nasional dari Masa ke Masa (National Examination from time to time). Retrieved July 2, 2015, from Okezone.com: <http://news.okezone.com/read/2014/12/30/65/1085698/ujian-n asional-dari-masa-ke-masa>
- [38]O'Keefe, T. P. (2012). A sequence labelling approach to quote attribution. *Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Language Learning*, (pp. 790-799).
- [39]Ofqual. (2014). *An Update on the Reforms Being Made to GCSEs*. London: Office of Qualifications and Examinations Regulation.
- [40]O'Keefe, T., Curran, J. R., Ashwell, P., Koprinska, I., & Honnibal, M. (2012). A Sequence Labelling Approach to Quote Attribution. *Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language*, (pp. 790-799). Jeju Island, Korea.
- [41]O'Keefe, T., Curran, J. R., Ashwell, P., & Konrinska, I. (2013). An Annotated Corpus of Quoted Opinions in News Articles. *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics*, (pp. 516-520). Sofia, Bulgaria.
- [42]Pang, B., & Lee, L. (2008). *Opinion Mining and Sentiment Analysis*. *Foundations and Trends in Information Retrieval*, 2 (1-2), 1-35.
- [43]Phelps, R. P. (2003). *Kill the messenger, The war on standardized testing*. New Brunswick,NJ: Transaction Publisher.
- [44]Polesel, J., Dulfer, N., & Turnbull, M. (2012). The experience of education: The impacts of high stakes testing on school students and their families. Sydney: Whitlam Insititute.
- [45]Scholz, T., & Conrad, S. (2012). Integrating Viewpoints into Newspaper Opinion Mining for a Media Response Analysis. *Proceedings of KONVENS*. Viena.
- [46]Sha, Y., Yan, J., & Cai, G. (2014). Detecting Public Sentiment Over PM2.5 Pollution Hazards through analysis of Chinese Microblog. In S. Hiltz, M. Pfaff, L. Plotnick, & P. Shih (Ed.), *Proceedings of the 11th International ISCRAM Conference*. University Park-Pennsylvania, USA.
- [47]Spar, A. (2012, May 20). Too much is riding on testing data that cannot be trusted. Retrieved June 20, 2015, from Orlando Sentinel: http://articles.orlandosentinel.com/2012-05-20/opinion/os-ed-f cat-test-score-andrew-spar-052012-20120518_1_fcat-reading-and-math-scores-from-one-year-teachers
- [48]Strauss, V. (2012, May 18). How standardized tests are affecting public schools. Retrieved July 10, 2015, from Washington Post: https://www.washingtonpost.com/blogs/answersheet/post/201 2/05/17/gIQABH1NXU_blog.html
- [49]Sukhum, K., Niruwat, S., & Haruechaiyasak, c. (2001). Opinion Detection in Thai Political News Columns Based on Subjectivity Analysis. *The 7th International Conference on Computing and Information Technology*..
- [50]Sutama, Narimo, S., & Samino. (2015). Management Of Curriculum 2013 Mathematic Learning Evaluation In Junior High School. *International Journal of Education*, 7 (3).
- [51]Suyitno. (2013, April 25). *Metamorfosis Ujian Nasional (Metamorphosis of National Examination)*. Retrieved June 10, 2015, from Kompas.com: <http://www.tempco.co/read/kolom/2013/04/25/696/Metamorfo sis-Ujian-Nasional>
- [52]Syahril, I., & Lesko, N. (2007). Standardized testing in Indonesian secondary education: An analysis on the impact of national exit exam (2005-2007). *Teacher College, Columbia University*. Retrieved April 6, 2015, from Scribd.com: <https://www.scribd.com/doc/293828706/Standardized-Testing -in-Indonesia>
- [53]Texas, S. o. (2013). S.B. No 1 general appropriations bill.

- Austin Texas: Legislative Reference Library of Texas.
- [54]Thompson, G. (2013). NAPLAN, MySchool and Accountability: Teacher perceptions of the effects of testing. *The International Education Journal: Comparative Perspectives*, 12 (2), 62-84.
- [55]Thompson, G., & Harbaugh, A. (2013). A preliminary analysis of teacher perceptions of the effect of NAPLAN on pedagogy and curriculum. *Australian Education Researcher*, 40 (3), 299-314.
- [56]Upe, I., & Wahyuddin. (2014). The Seed of Educational System in Indonesia Based on the UN (National Examination) Results. *International Symposium on Advances in Technology Education*. Nanyang Polytechnic, Singapore.
- [57]Vilares, D., Thelwall, M., & Alonso, M. (2015). The megaphone of the people? Spanish SentiStrength for real-time analysis of political tweets. *Journal of Information Science* , 1-16.
- [58]WelcomeTrust. (2011). *Primary Science Survey Report*. London: Welcome trust.
- [59]Wiebe,J., Wilson,T., Bruce,R., Bell, M., & Martin, M. (2004). Learning Subjective Language. *Computational Linguistics*, 30 (3), 277-308.