

Perceptions of Government External Auditors on The Effectiveness of Red Flags in Detecting Fraud on Procurement of Goods and Services in West Java

Hengki

Widyatama University
hengki.3336@widyatama.ac.id

Iman Budiman

Widyatama University
iman.budiman@widyatama.ac.id

Silviana

Widyatama University
silviana.msi@widyatama.ac.id

Abstract

This study aimed to determine which red flags are considered effective for detecting fraud by auditors and to determine differences in perceptions among auditors with different demographic factors (gender, certifications: certified fraud examiner (CFE)/ certified fraud auditor (CFrA), tenure, position, and fraud-related experience) regarding the effectiveness of red flags on the procurement of goods and services. Respondents were the government external auditors who work at the Audit Board of the Republic of Indonesia (BPK RI), West Java Province Representatives by obtaining data through questionnaires. Perceptions about the effectiveness of red flags were measured by an interval scale of 41 red flags statements adopted from the Decree of the BPK RI concerning Technical Instructions for Inspection of Procurement of Goods and Services and Inspection of Infrastructure Expenditures in 2019. Data analysis was performed using the independent sample t-test and ANOVA test. The results showed that the most effective red flags perceived by the respondent for detecting fraud in the procurement of goods and services were "Owner price (HPS) prepared by prospective participants/together with the officer who signed the agreement (PPK)" ($mean = 4.240$), "Ad hoc team selection and/or PPK who provide detailed information on HPS to prospective bidders" ($mean = 4.170$), and "There is a fictitious project" ($mean = 4.140$). In general, there were no differences in the perception among auditors with different demographic factors except for the difference in tenure ($p > 0.05$). It can be concluded that the most effective red flags perceived to detect fraud on the procurement of goods and services was "Owner price (HPS) prepared by prospective participants/ together with the officer who signed the agreement (PPK)" and there was a difference in the perceptions among auditors based on the tenure.

Keywords

Auditor perceptions, demographic factors, red flags, fraud, procurement of goods and services.

To cite this article: Hengki, Budiman, I., Silviana (2021) Perceptions of Government External Auditors on The Effectiveness of Red Flags in Detecting Fraud on Procurement of Goods and Services in West Java. *Review of International Geographical Education (RIGEO)*, 11(8), 2551-2563. doi: 10.48047/rigeo.11.08.237

Submitted: 09-10-2020 • **Revised:** 11-12-2020 • **Accepted:** 13-02-2021

Introduction

Corruption in the government body has made Indonesia one of the most corrupt countries in Southeast Asia. Based on the Data Corruption Perceptions Index of 2020, out of 183 countries surveyed, Indonesia ranks 102 with the GPA of 37, down 13 rankings from 2018 (CPI, 2020) This condition shows that the management of state finances is still far from achieving effectiveness and efficiency because there are still many leaks through corrupt practices.

The public questioned the true meaning of Unqualified (Utomo, Kriek, Labbé, Conroy, & Fumagalli, 2014) opinion in a government agency issued by the Indonesian Supreme Audit Agency (BPK), why the government agencies with the predicate of Unqualified apparently did not immediately clear of fraud. Corruption cases in entities that are awarded as WTP have eroded public trust in BPK.

In its press release on September 17, 2019, BPK found 14,965 problems worth IDR 10.35 trillion in the examination during the first semester of 2019. The number includes 7,236 problems with the weakness of the institution's internal control system, 7,636 problems of non-compliance with the provisions of laws and regulations worth IDR 9.68 trillion, and 93 problems of ineffectiveness, inefficiency, and ineffectiveness worth IDR 676.81 billion (Arens, Elder, & Beasley, 2015)

Although the BPK has found various cases of violations of provisions in the process of procurement of goods and services, this is not enough to reduce the officials' bad intentions to commit fraud. The number of public complaints related to acts of fraud against public officials shows that fraud is still rampant in Indonesia. On the other hand, auditor also have limitations to detect all acts of fraud.

Stated that forensic accounting is very much needed in detecting fraud (Bierstaker, Brody, & Pacini, 2006) Rezaee suggested that auditors should use more effective and objective audit procedures and related standards to increase audit effectiveness (Rezaee, Crumbley, & Elmore, 2004) However, Hassink et al revealed that auditors are sometimes unable to uncover fraud due to time constraints, conflicts of interest, and non-compliance with standards (Hassink, Meuwissen, & Bollen, 2010). Therefore, a new effective way is required to help the auditor perform their duties properly.

Statement of Auditing Standard (SAS) No.99 AICPA identify the risk factors that can be used by auditors in conducting audits in the business sector (Arens et al., 2015) These standards are in line with the Professional Standards of Certified Public Accountants (SPAP) as contained in Auditing Standard (SA) 240 regarding the auditor's responsibilities regarding fraud in an audit of financial statements that include fraud factor in financial statements.

The auditing standards and guidelines contain a list of red flags that can be used by auditors to detect fraud during audit assignments. However, the existing audit standards and guidelines do not provide guidance to measure the effectiveness of each red flags item, which can lead to different perceptions of each auditor on the effectiveness of red flags in detecting fraud. A study identified states that red flags have a positive effect on the ability of auditors to detect fraud (Arens et al., 2015)

Several studies showed different results regarding which red flags are considered the most important for detecting fraud. Differences in the auditor's personal characteristics can lead to different perceptions (Robbins & Judge, 2008) A study done by Novitasari and Rustriarini's pointed out that the experience of the auditors affect on the auditor's perception of red flags (Rezaee et al., 2004). While research done by Utomo found that differences in the demographics do not lead to differences in the perception of the auditor of the red flags (Utomo et al., 2014)

Evidence related to the perception of the effectiveness of red flags showed inconsistent results, while the auditors' demographic factors showed varying results. Thus, it is important to conduct further research, especially in the context of public sector, especially on the procurement of goods and services in the government.

This study aimed to answer the following research questions:

- a. Which red flags are considered effective for detecting fraud in the procurement of goods and services?
- b. Are there any differences in the perceptions among auditors who have different demographic factors (gender, CFE/CFrA certification, tenure, position, and fraud-related experience regarding the effectiveness of red flags in the procurement of goods and services?

Literature review

a. Perception

The Big Indonesian Dictionary (KBBI) defines perception as "a direct response (acceptance) of something, or the process of someone knowing some things through his five senses." (Kurt & Hacıoglu, 2010) Apostolou et al found that auditor demographic factors, such as type of public accountant firm, tenure, auditor position, and industry specialization, affect auditor perceptions in assessing the effectiveness of red flags (Apostolou, Hassell, Webber, & Sumners, 2001) Siegel and Ramanauskas-Marconi stated that the auditor's judgment is influenced by the perception of a situation (Siegel & Ramanauskas-Marconi, 1989)

b. Procurement of Goods and Services

Presidential Regulation Number 16 in 2018 concerning Government Procurement of Goods and Services defines the procurement of goods and services as a process of activities to obtain goods and services by ministries/agencies/regional work units/institutions using the state budget, either central budget (APBN) or regional (APBD) whose process starts from planning the requirements to the completion of all activities in order to obtain goods and services (Publik, 2011)

c. Independent Auditor

Boynton et al defines auditors as professionals assigned to audit economic activities and events for individuals and legal entities (Boynton, Johnson, & Kell, 2001). State Financial Audit Standards (SPKN) defines an auditor as an officer who examining the management and responsibilities of state finances for and on behalf of BPK. The independent auditor in this study is the government's external auditor who works at the BPK Representative for West Java Province.

d. Fraud

SPKN defines fraud as "The deed containing the element of intention; intention; self-benefit or for someone else; deception; concealment or embezzlement; and abuse of trust to gain illegitimate advantage in a form of money, goods/property, services, and do not pay for the services, which are done by one individual or more from the official parties responsible for governance, employees, or a third party" (Smith, Omar, Idris, & Baharuddin, 2005). Fraud or fraudulent acts of fraud was intentionally done to take property of others or a deliberate misstatement in the financial statements (Arens et al., 2015) The concept of the fraud triangle was introduced by Donald Cressey which emphasizes three conditions that cause fraud: pressure (incentive), opportunity, and attitude to rationalize action (rationalization/attitude) (Herusetya, 2020)

e. Red Flags

Singleton defines the red flags as varying conditions that include things such as accounting anomalies, transactions or events that cannot be explained, unusual elements of a transaction and change a person's behavior (Singleton, 2010). A study done by Smith et al identified the red flags considered most important by auditors, and explored whether auditor demographic factors (gender, auditor experience, auditor tenure, type of public accountants, and fraud-related experience) might have an impact on auditors perceptions regarding the importance of red flags in Malaysia (Smith et al., 2005). In this research lead to the conclusion that red flags are important for the auditor in an effort to detect fraud.

Research Hypothesis

Differences in demographic characteristics or demographic factors may lead to different perceptions. The hypotheses of this research are:

H₁ : There are differences in the perception of auditors who have gender differences regarding

the effectiveness of red flags in the procurement of goods and services.

H₂ : There is a difference in the perception between auditors who have certification CFE/CFrA with those who do not have CFE/CFrA certification regarding the effectiveness of red flags in the procurement of goods and services.

H₃ : There are differences in the perception of auditors who have different tenures regarding the effectiveness of red flags in the procurement of goods and services.

H₄ : There are differences in the perceptions of auditors who have different positions regarding the effectiveness of red flags in the procurement of goods and services.

H₅ : There are differences in the perceptions of auditors who have different experiences related to fraud regarding the effectiveness of red flags in the procurement of goods and services.

Method

a. Data Types and Sources

Primary data used in this study were obtained from a direct survey by sending questionnaires to respondents. The survey was designed to capture the concept of definitions that reflect the strength of attitudes, perceptions, views, and opinions from government's external auditor who works at the BPK Representative of West Java Province.

The questionnaire consists of two parts:

1. The first part is the respondents' demographic information;
2. The second part is questions regarding the auditors' fraud-related experiences during their auditing tasks, and also collects a number of respondents perceptions about the effectiveness of the red flags adopted from concerning Technical Guidelines for Auditing Procurement of goods and services and The Inspection Policy of Infrastructure Expenditure 2019 which are grouped into three steps: (1) Preparation; (2) Selection of goods and services providers and the signing of the contract; and (3) Implementation and Reporting steps (Boynnton et al., 2001)

b. Population and Sampling Procedure

The population is the government's external auditor who works at the BPK Representative province of West Java.

c. Research Variable

1. Perception of Red Flags Effectiveness Level

Perceptions of the effectiveness of red flags were measured using 41 statements. Red flags as the research instruments were adopted from BPK RI Decree No. 9/K/I-XIII.2/10/2009 regarding Technical Guidelines for Inspection of Procurement of Goods and Services, and The Inspection Policy of Infrastructure Expenditure 2019. The technical guidelines and inspection policies provide red flags, which are marked as tipping points. Details of research variable presented in Appendix 1.

The answers to each question are measured using six interval scales to determine the auditor's tendency to assess the effectiveness of red flags, namely: (1) very ineffective; (2) not effective; (3) less effective; (4) quite effective; (5) effective; and (6) very effective.

2. Auditor Demographic Factors

Auditor demographic factors include gender, certification, tenure, position, and fraud-related experience. This variable will be tested for differences analysis in auditor perceptions of the effectiveness of red flags in detecting fraud in each category. Demographic factors are described on Figure 1.

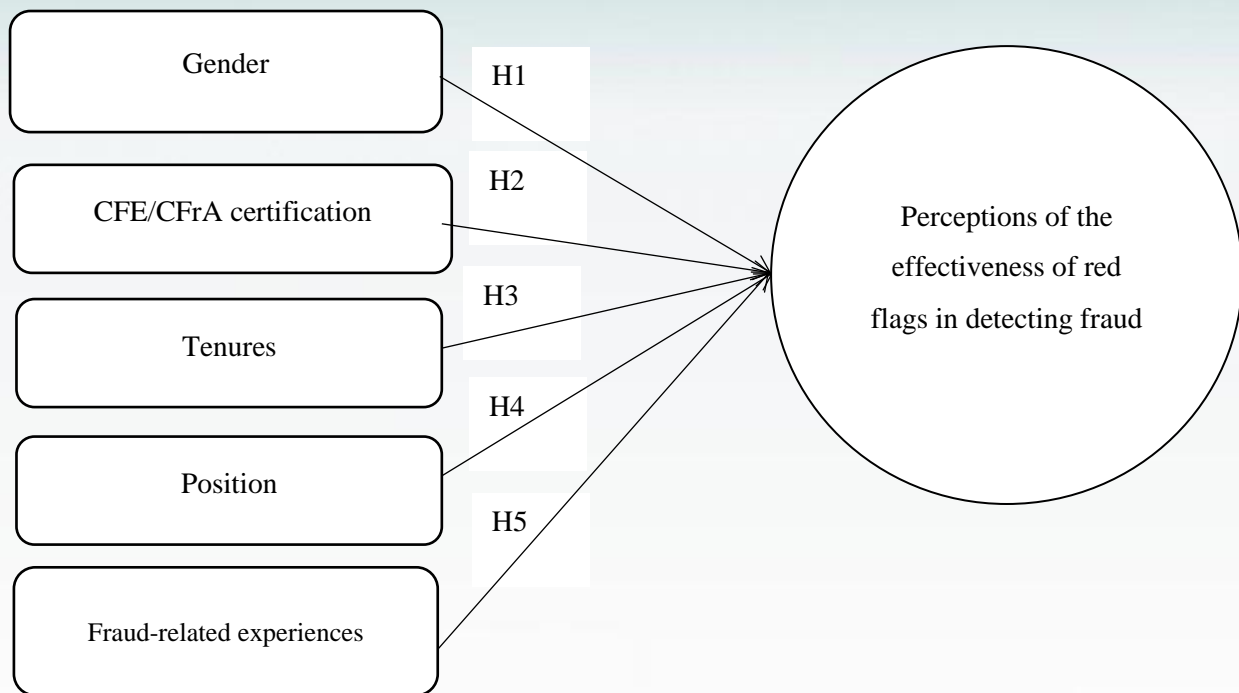


Figure 1. Research Model Framework

d. Data Quality Test

1. Validity Test

Validity test was done to determine the precision and accuracy of a measuring instrument to carry out its measuring function. The validity test was done using the Pearson Product Moment correlation from the SPSS 19®.

2. Reliability Test

Measurement was done by comparing one question to another or measuring the correlation between the answers to the questions. This test is carried out using Cronbach Alpha statistical test from SPSS 19®. A variable is considered reliable if the Cronbach Alpha value > 0.60 (Hair et al., 2006).

e. Data analysis technique

1. Descriptive statistics

We used descriptive statistics to describe the data by calculating the mean (m) and standard deviation (SD), and the frequency or presentation from the respondent. This analysis is needed to describe perception of the overall red flags effectiveness.

The data was analysed according to the following steps:

- a) Calculate the mean value for each red flags item,
- b) Calculate the overall mean value of the red flags,
- c) Compare the mean value of each item with the overall mean value;
- d) Classify the perception of effectiveness level of red flags into three categories: "less effective", "fairly effective", and "effective" with reference to the mean value; and
- e) Draw conclusions by:
 - 1) If the mean value of a red flags item is significantly higher than the overall mean value, then the item is considered an "effective" red flags group in detecting fraud;
 - 2) If the mean value of a red flags item is not significantly different from the overall mean value, then the item is considered a "quite effective" red flags group in detecting fraud; and
 - 3) If the mean value of an item's red flags is significantly lower than the overall mean value, then the item is considered a "less effective" red flags group in detecting fraud.

2. Factor Analysis

This study adopted 36 items of red flags from BPK RI Decree No. 9/K/I-XIII.2/10/2009 concerning the Technical Guidelines for Inspection of Procurement of Goods and Services, and 18 items of red flags from The Inspection Policy of Infrastructure Expenditure 2019. The 54 red flags items were further simplified into 41 red flags items and grouped into three major groups: the preparation stage, the selection of goods and services providers and the signing of contracts, as well as implementation and reporting. The grouping is intended to summarize the 41 red flags items into new dimensions or variables (factors) to provide a concise description of the red flags items and as a main fraud indicator for the auditor.

The stages of factor analysis carried out are as follows:

a) Kaiser-Meyer-Olkin (KMO) and Bartlett's test

To run the factor analysis, the results of the KMO test must be greater than 0.5. Bartlett's test was used to observe whether the correlation matrix is the identity matrix. The result must be significant, which means that the correlation matrix is not an identity matrix so that factor analysis can be used.

b) Anti-Image Metrics

This test is conducted to see what variables are suitable for factor analysis. A variable is considered feasible if the MSA (Measures of Sampling Adequacy) value is > 0.5 .

c) Factor Extraction

Determination of the number of factors using the Determination Based on Eigen Value approach. In this approach, only factors that have an eigen value > 1 will be used for further research.

3. Independent Sample t-test

This study used independent sample t-test to compare the average of two sample groups that are not related to each other (Ghozali, 2011) with demographic factors such as gender, certification, and fraud-related experiences.

4. Analysis of Variance (ANOVA)

This analysis was conducted to determine the main effect and interaction effect of the independent categorical variables on the metric dependent variable (Ghozali, 2011). This test is conducted to determine the comparison of the average auditor's perception of red flags in the procurement of goods/services based on tenure and position.

Analysis and Discussion

a. Characteristics of Respondents

1. Sample and Questionnaire Return Rate

This sample was auditors who work at the BPK Representative of West Java Province. Questionnaires were distributed to respondents via electronic mail. A total of 132 questionnaires were sent via electronic mail and 100 responses were obtained (75.76%).

2. Respondents Overview

The general description of the respondents is depicted in the demographic factors of each respondent (table 1).

Table 1
Overview of Respondents

	Sample	Frequency	Percentage
Gender	Man	72	72.00 %
	Woman	28	28.00 %
CFE/CFrA Certifications	Yes	5	5.00 %
	No	95	95.00 %
Working period (years)	0-5	1	1.00 %
	6-10	20	20.00 %
	>10	79	79.00 %
Position	First Expert Examiner	28	28.00 %
	Young Expert Examiner	56	56.00 %
	Associate Examiner	16	16.00 %
	Main Expert Examiner	0	0.00 %
Fraud-related experience	Yes	64	64.00 %
	No	36	36.00 %

b. Data Quality Test

The validity test showed that the calculated r value on all question items was positive and greater than the standard r value, which means all questions on the questionnaire are valid. Additionally, the reliability test showed the Cronbach's Alpha was 0.9929, which indicated that the variable is reliable.

c. Data Analysis Results

1. Descriptive statistics

The descriptive statistics test indicated that the mean value of all red flags items was 3.712. It also showed that the top three most effective red flags in detecting fraud perceived by the auditors were "HPS prepared by prospective participants/ together with PPK", "Ad hoc team selection and/or PPK who provide detailed information on HPS to prospective bidders", and "there is a fictitious job" with the mean value were 4.240, 4.170, and 4.140, respectively.

The classification of fraud detection effectiveness levels from 41 red flags items showed that there were 16 red flags items that are perceived effective in detecting fraud, 11 red flags items were perceived to be quite effective, and the remaining 14 red flags items were perceived to be less effective in detecting fraud in procurement of goods and services.

2. Factor Analysis

Factor analysis is intended to summarize the information content of a large number of variables into a number of smaller factors.

a) Kaiser-Meyer-Olkin and Bartlett's test

The results of the Kaiser-Meyer-Olkin (KMO) test and Bartlett's test between variables on the red flags item in this study is represented in Table 2.

Table 2
KMO and Bartlett's Test Result

Kaiser-Meyer-Olkin Measure of Sampling Adequacy		0.966
Bartlett's Test of Sphericity	Approx. Chi-Square	6903.997
	Df	820
	Sig.	0.000

Table 2 shows that the results of the KMO test was 0.966. This value is greater than the minimum required limit, which is 0.5, while the value of Bartlett's test showed a significance value of 0.000

which means that the correlation matrix is not an identity matrix, which indicates that the factor analysis can be used for further assessment.

b) Anti-Image Metrics

Anti-Image Metrics are used to assess which variables are eligible to be used in the factor analysis. All variables have MSA values greater than the minimum requirement of 0.5. This means that all red flags variables are eligible to measure the variables for further analysis.

c) Factor Extraction

In this study, factor extraction was carried out using a determination based on eigen value approach, which means that the factor that will be used for further testing is a factor that has an eigen value >1.

Table 3

Total Variance Explained

Component	Initial Eigen values		
	Total	% of variance	Cumulative (%)
1	32.264	78.694	78.694
2	1.378	3.362	82.056
3	1.039	2.533	84.589

The results showed that there were three factors that have an eigen value >1. The variance value showed the contribution of the new factor to the overall red flags factor. Factor 1 (F1) has a variance value of 78.694%, which indicates that factor 1 contributes 78.694% to the red flags factors. Cumulative value of the three factors (F1, F2, and F3) formed is 84.589%, which means that the three factors are able to explain 84.589% of the 41 red flags variables. The remaining 15.411% is explained by other factors.

3. Independent Sample t-test

a) T-Test Based on Auditor Gender

The results showed that the average perception of male respondents in assessing the effectiveness of red flags is greater than the perception of female respondents. This shows that there is a difference in the average perception of male auditors on the effectiveness of red flags when compared to female auditors.

Table 4

Average Auditor Perception by Gender

Items	Gender	mean	Std. Deviation
F1	Man	86.14	29.493
	Woman	70.25	37.458
F2	Man	36.64	13.366
	Woman	31.57	17.048
F3	Man	36.57	11.545
	Woman	32.00	13.725

The next stage of analysis is to test whether the assumptions of the population variance of the two samples are the same (equal variance assumed) or different (equal variance not assumed) by looking at the Lavender's test value. The null hypothesis in this test is that the population perception of the effectiveness of red flags between male and female auditors is the same. If probability >0.05, then the null hypothesis cannot be rejected, which means that the variances are the same. Lavene's test showed that the F value on all items is significant ($p < 0.05$), which means that it rejects the null hypothesis which states that the variance is the same. This means that the assumption of the data having the same variance is not comply. Thus, the analysis of the t-test difference must use equal variance not assumed.

Table 5
T-Test Based on Auditor Gender

Items	Levene's Test for Equality of Variances		t-test for Equality of Means			
	F	Sig.	t	df	Sig. (2- tailed)	
F1	Equal variances assumed	11.109	0.001	2.237	98	0.028
	Equal variances not assumed			2.015	40.689	0.051
F2	Equal variances assumed	8.252	0.005	1.572	98	0.119
	Equal variances not assumed			1.413	40.571	0.165
F3	Equal variances assumed	4.069	0.046	1.684	98	0.095
	Equal variances not assumed			1.560	42.675	0.126

The independent sample t-test was conducted by determining the null hypothesis: there was no difference in the average perception of male and female auditors. The null hypothesis is accepted if the significance probability is greater than 0.05 ($p\text{-value} > 0.05$). The t-value for F1 is 2.015 with a significance probability ($p\text{-value}$) of 0.051. The $p\text{-value}$ was > 0.05 , so it cannot reject the null hypothesis. This means there is no difference in the average perception of male and female auditors. The t-value for F2 was 1.413 with probability significance was 0.165 ($p\text{-value} > 0.05$), this means there is no difference in the average perception between male and female auditors. The t-value for F3 was 1.560 with a probability significance of 0.126 ($p\text{-value} > 0.05$) which means that there is no difference in the average perception between male and female auditors. The results of the t-test by gender for items F1, F2 and F3 indicated that there is no difference in the average perception of male and female auditors.

b) T-Test Based on Certification Ownership

The t-test based on auditors's certification, specifically the CFE/CFrA certification. Auditors who have CFE/CFrA certification are assumed to have a better level of understanding regarding fraud detection compared to auditors who do not have the certification.

Table 6
Average Auditor Perception Based on The Ownership of CFE/CFrA Certification

Items	Certification	Mean	Std. Deviation
F1	Non-CFE/CFrA	81.03	33.297
	CFE/CFrA	88.33	23.754
F2	Non-CFE/CFrA	81.03	33.297
	CFE/CFrA	36.89	11.407
F3	Non-CFE/CFrA	35.04	12.613
	CFE/CFrA	37.78	8.599

Table 6 showed that the average perception of auditors who do not have CFE/ CfrA certification tends to be greater than the one who do not.

Table 7
T-test Results Based on The Possession of Certification

Items	Levene's Test for Equality of Variances		t-test for Equality of Means			
	F	Sig.	T	Df	Sig. (2- tailed)	
F1	Equal variances assumed	3.728	0.056	0.640	98	0.523
	Equal variances not assumed			0.844	11.373	0.416
F2	Equal variances assumed	1.109	0.295	0.358	98	0.721
	Equal variances not assumed			0.446	10.900	0.664
F3	Equal variances assumed	2.428	0.122	0.634	98	0.527
	Equal variances not assumed			0.866	11.720	0.404

The Levene's test in table 7 showed that the F values on items F1, F2 and F3 are not significant

($p > 0.05$), which means they cannot reject the null hypothesis which states that the variance is the same. This means that the assumption of the data having the same variance is accepted. Thus, the analysis of the t-test difference test must use equal variance assumed.

The t value for F1 is 0.640 with a significance of 0.523. The p value > 0.05 , so it cannot reject the null hypothesis which states that there is no difference in the average perception of the effectiveness of red flags between auditors who have CFE/CfrA certification and auditors who do not have CFE/CfrA certification. T-test for F2 item using equal variance not assumed, the value of the t was 0.358 with a significance of 0.721 (p-value > 0.05) for F2, which means that there is no difference between the average perception of auditor CFE/CfrA and auditor non-CFE/CfrA. In an item F3, the value of the t was at 0.634 with a significance probability of 0.527 (p value > 0.05), which means that to item F3 there is no average difference in the perceptions between CFE/CfrA and non-CFE/CfrA auditors.

On the same fashion, the t-test results showed that for items F1, F2, and F3 there is no difference in the average perception of CFE/CfrA non-CFE/CfrA auditors. Based on these results, there is no difference in the average perception between auditors who have CFE/CfrA certification and not regarding the effectiveness of red flags in detecting fraud in the procurement of goods and services.

c) T-Test based on fraud-related experience.

A t-test based on fraud-related experience was conducted to analyse whether there is a difference in the average perception between auditors who have found/disclosed fraud and auditors who have never found fraud, as shown in table 8.

Table 8
Auditor Perception Based on Fraud-related Experience

Items	Auditor Experience	Mean	Std. Deviation
F1	Did not find fraud	80.58	32.430
	Have you ever found fraud?	82.31	32.818
F2	Did not find fraud	34.92	15.030
	Have you ever found fraud?	35.39	14.440
F3	Did not find fraud	36.25	11.970
	Have you ever found fraud?	34.75	12.538

The results on table 8 showed that there are variations in the average value of the respondent's perceptions. It showed that there is a difference in the average perception of the two groups of auditors. Furthermore, t-test was done to determine whether the assumption of data variance is comply by looking at the results of lavender's test. Lavene's test showed that the F value for all items F1, F2, and F3, was not significant ($p > 0.05$), which mean it cannot reject the null hypothesis, which states that the variance is the same. It indicated that the assumption of the data having the same variance is comple. Thus, the analysis of the t-test difference test used equal variance assumed.

Table 9
T-test Results Based on Fraud-related Experience

Items	Levene's Test for Equality of Variances		t-test for Equality of Means			
	F	Sig.	T	df	Sig. (2- tailed)	
F1	Equal variances assumed	0.022	0.881	0.254	98	0.800
	Equal variances not assumed			0.255	73.404	0.800
F2	Equal variances assumed	0.371	0.544	0.155	98	0.877
	Equal variances not assumed			0.154	70.259	0.878
F3	Equal variances assumed	0.069	0.793	-0.584	98	0.561
	Equal variances not assumed			-0591	75.547	0.556

Table 9 showed that the t-value for F1 was 0.254 with a probability significance of 0.800 (p-value > 0.05). It means the null hypothesis-which states that there is no difference in the average

perception of the effectiveness of red flags between groups of auditors who have and have not experienced of disclosed fraud-cannot be rejected. The t-test for item F2 showed a t-value of 0.155 with a significance of 0.877 (p value > 0.05). This means there is no difference in the average perception of the two groups of auditors (F2). The t-value on the F3 item was -0.584 with a probability significance of 0.561 (p -value > 0.05), meaning that is no difference in the average perception.

The t-test results for items F1, F2, and F3, all showed there is no difference in the average perception between auditors who have and have not experienced of disclosing fraud.

4. Analysis of Variance (ANOVA)

a) ANOVA Test Based on Service Period

Initial testing was conducted to test the assumption that each category of independent variables has the same variance. The assumption of data having the same variance is obtained if the significance level is greater than 0.05.

Table 10

Homogeneity of Variances Test Based on Service Period

Items	Levene Statistics	df1	df2	Sig.
F1	0.527	1	97	0.470
F2	0.005 ^a	1	97	0.944
F3	0.296 ^a	1	97	0.588

The results of lavender's test showed that the F-value for each item was not significant at 0.05 (p value > 0.05), which means it cannot reject the null hypothesis which states that the variance is the same. This means that the ANOVA assumption is accepted that the variance is the same.

The following step was to test the level of significance for each item. If the significance level is less than 0.05 (p -value < 0.05), then there is a significant difference in perceptions of the effectiveness of red flags in each working group.

ANOVA test results are presented in Table 4.11. It indicated that the F value is 5.512 (F1) with a significance of 0.005 (p value < 0.05), meaning there were differences in perceptions of red flags F1 between groups of auditors. For F2, it showed an F value of 5.303 with a significance of 0.007 (p -value < 0.05), so it means that there are differences in perceptions of F2 between groups of auditors. The results of the ANOVA test for F3 gave an F value of 4.030 with a significance value of 0.021 (p -value < 0.05) which also indicated that there were differences in perceptions of red flags between groups of auditors.

Table 11

ANOVA Test Results Based on Service Period

	Items	Sum of Squares	df	Mean Square	F	Sig.
F1	Between Groups	10688.232	2	5344.116	5.512	0.005
	Within Groups	94043.158	97	969.517		
	Total	104731.390	99			
F2	Between Groups	2074.676	2	1037.338	5.303	0.007
	Within Groups	18974.484	97	195.613		
	Total	21049.160	99			
F3	Between Groups	1148.551	2	574.276	4.030	0.021
	Within Groups	13822.039	97	142.495		
	Total	14970.590	99			

Based on the ANOVA test, it can be denoted that there are differences in the average perception of the effectiveness of red flags in detecting fraud in the procurement of goods/services between groups of auditors with different tenures.

b). Anova test based on position

The results of lavender's test showed that the F value for each item is not significant at 0.05 (p value < 0.05), which means it cannot reject the null hypothesis which states that the variance is the same. This means that the ANOVA assumption is accepted (the variance is the same). More details on Lavender's test are displayed on Table 12.

Table 12

Homogeneity of Variances Test
Base on Auditor Position

Items	Levene Statistics	df1	df2	Sig.
F1	1.755	2	97	0.178
F2	0.699	2	97	0.499
F3	0.742	2	97	0.479

The next step is to test the level of significance for each item. If the significance level is less than 0.05 (p value < 0.05) means that there are significant differences in perceptions of the effectiveness of red flags in each group.

Table 13

ANOVA Test Based on Auditor Position

Items		Sum of Squares	df	Mean Square	F	Sig.
F1	Between Groups	1127.506	2	563.753	0.528	0.592
	Within Groups	103603.884	97	1.068.081		
	Total	104731.390	99			
F2	Between Groups	231.526	2	115.763	0.539	0.585
	Within Groups	20817.634	97	214.615		
	Total	21049.160	99			
F3	Between Groups	134.447	2	67.224	0.440	0.646
	Within Groups	14836.143	97	152.950		
	Total	14970.590	99			

The ANOVA test showed that for the F1 item, the F value is 0.528 with a significance of 0.592 (p value > 0.05). This means no difference in the perception of red flags F1 between groups of auditor positions. Similarly, the F2 dan F3 item test showed an F value of 0.539 and 0.440 with a significance of 0.585 (p value > 0.05) and 0.646 (p value > 0.05), respectively, which means there is also no difference in the perception of F2 and F3 between groups of auditor positions.

Discussion

The classification of fraud detection effectiveness levels from 41 red flags items showed that there were 16 red flags items perceived as effective in detecting fraud, 11 red flags items were perceived to be quite effective, and the remaining 14 red flags items were perceived to be less effective in detecting fraud in procurement goods and services. The statistical analysis showed that the red flag items that were perceived to be the most effective are "HPS prepared by prospective participants/together with PPK," "Working group Selection and/or PPK provide detailed information on HPS to prospective bidders", and "There is a fictitious job", respectively. The factor analysis grouped the 41 red flags items into three new red flags factors. The first factor was "the preparation stage of the procurement of goods/services and the selection of providers of goods/services is not in accordance with the provisions". This factor comes from red flags items related to problems at the preparation stage for the procurement of goods/services and the selection stage for the provision of goods/services that are not in accordance with the provisions, both from the committee side, HPS preparation, as well as the tender winner selection process. The second factor is "implementation and reporting not in accordance with the provisions", which comes from red flags items related to problems at the stage of implementation and reporting work, both related to the quality and quantity of the project, as well as the preparation of final reports and payments. Third factors was "the process of selecting suppliers of goods/services, and the signing of the agreement that is not in accordance with provisions" relating to the issue of documents on the tender process and document of agreement. "The preparation stage for the procurement of goods/services and the selection of goods/services

providers is not in accordance with the provisions" was a red flag statement which perceived to be the most effective for detecting fraud in the procurement of goods and services.

T-test results based on the type of gender showed that there is no average difference between the male and female auditor's perception which makes H_1 is not accepted.

The t-test based on the CFE/CfrA certification indicated that there is no difference in the average perception of CFE/CfrA auditors and non-CFE/ CfrA auditors.

The results of the ANOVA test based on the auditor's tenure indicated that there is a difference in the perception of red flags between the auditor's tenure groups. Based on the empirical test results, the H_3 which states that there are differences in the perception of the auditor with different working duration on the effectiveness of red flags in procurement of goods and services is accepted.

The ANOVA test was done based on the group of auditor positions, and it showed that H_4 which states that there are differences in the perception of auditors who have different positions regarding the effectiveness of red flags in the procurement of goods/services is not accepted. Moreover, the t-test was done based on the experience of the auditors related to fraud, and it showed that for all values of F1 ($t=0.254$; $p\text{-value} = 0.800$), F2 ($t = 0.155$; $p\text{-value} = 0.877$), and F3 ($t = -0.584$; $p\text{-value} = 0.561$) indicated no difference of perception between the auditor who had and had not experience in revealing fraud, which means the H_5 is not accepted.

Conclusions and suggestions

Conclusion

These are several conclusions drawn from this study:

1. The red flags that are perceived to be the most effective for detecting fraud in the procurement of goods/services are "HPS compiled by prospective participants/together with PPK, for example: similarity of properties on softcopy of HPS and budget for bidding document files, "Working group selection and/or PPK provide detailed information on HPS to potential bidder", and "there is a fictitious job". The classification of fraud detection effectiveness levels from 41 red flags items showed that there were 16 red flags items were perceived as effective in detecting fraud, 11 red flags items were perceived to be quite effective, and the remaining 14 red flags items were perceived to be less effective in detecting fraud in procurement goods and services;
2. The factors analysis indicated that 41 red flags were grouped into a new three red flags factors, which were "the preparation stage of the procurement of goods and services and selection of goods and services not in accordance with the provisions", "implementation and reporting does not match provisions", and "the process of selecting providers of goods and services, and signing contracts that do not comply with the provisions".
3. Among the new three red flag factors, one that perceived as the most effective way to detect fraud was "preparation stage of the procurement of goods and services and selection of goods and services not in accordance with the provisions";
4. The empirical test showed that there were differences in the perception of auditors who have different tenures regarding the effectiveness of red flags in the procurement of goods and services (significant at $p\text{ value} < 0.05$) for the three factors F1, F2, and F3. Moreover, the empirical test showed that there is no difference in the average perception between auditors with demographic factors based on gender, ownership of CFE /CfrA certification, fraud-related experience, and position regarding the effectiveness of red flags in detecting fraud in the procurement of goods and services;

Suggestion

Suggestions for further research, includes:

1. Wider range of research using larger sample which covers all elements of the population, such as government internal auditors.
2. Involves the electronic data of procurement of goods/services (e-procurement). This will present the most updated problems exist in the procurement process.

References

- Apostolou, B. A., Hassell, J. M., Webber, S. A., & Sumners, G. E. (2001). The relative importance of management fraud risk factors. *Behavioral Research in Accounting*, 13(1), 1-24. doi:<https://doi.org/10.2308/brja.2001.13.1.1>
- Arens, A. A., Elder, R. J., & Beasley, M. S. (2015). *Auditing & Jasa Assurance*. Jakarta: Erlangga.
- Bierstaker, J. L., Brody, R. G., & Pacini, C. (2006). Accountants' perceptions regarding fraud detection and prevention methods. *Managerial Auditing Journal*. doi:<https://doi.org/10.1108/02686900610667283>
- Boynton, W. C., Johnson, R. N., & Kell, W. G. (2001). *Modern auditing*: J. Wiley & Sons.
- Ghozali, M. L. (2011). *Percikan Samudra Hikmah: Syarah Hikam Ibnu Atho'llah As-Sakandari*: Kencana.
- Hassink, H., Meuwissen, R., & Bollen, L. (2010). Fraud detection, redress and reporting by auditors. *Managerial Auditing Journal*. doi:<https://doi.org/10.1108/02686901011080044>
- Herusetya, A. (2020). Audit Quality Of Big Four Firms: Evidence From Client's Business Strategy. *Jurnal Reviu Akuntansi dan Keuangan*, 10(3), 430-443.
- Kurt, G., & Hacıoglu, G. (2010). Ethics as a customer perceived value driver in the context of online retailing. *African Journal of Business Management*, 4(5), 672-677.
- Publik, S. P. A. (2011). Institut Akuntan Publik Indonesia. Jakarta: Salemba Empat.
- Rezaee, Z., Crumbley, D. L., & Elmore, R. C. (2004). Forensic accounting education: A survey of academicians and practitioners. *Advances in Accounting Education, Forthcoming*.
- Robbins, S. P., & Judge, T. A. (2008). *Organizational behavior* 15th edition. US: Prentice Hall.
- Siegel, G., & Ramanauskas-Marconi, H. (1989). *Behavioral accounting*: Thomson South-Western.
- Singleton, T. W. (2010). *Fraud auditing and forensic accounting* (Vol. 11): John Wiley & Sons.
- Smith, M., Omar, N. H., Idris, S. I. Z. S., & Baharuddin, I. (2005). Auditors' perception of fraud risk indicators: Malaysian evidence. *Managerial Auditing Journal*. doi:<https://doi.org/10.1108/02686900510570713>
- Utomo, D., Kriek, M., Labbé, I., Conroy, C., & Fumagalli, M. (2014). Simultaneous Modeling of the Stellar and Dust Emission in Distant Galaxies: Implications for Star Formation Rate Measurements. *The Astrophysical Journal Letters*, 783(2), L30. doi:<https://doi.org/10.1088/2041-8205/783/2/L30>