

Cek Plagiarisme Publikasian

by Lufi Mursita

Submission date: 09-Sep-2020 08:00PM (UTC+0700)

Submission ID: 1382766362

File name: M_-_Penelitian_Artikel_Jurnal_Nasional_Terakreditasi_Sinta_2.pdf (425.49K)

Word count: 10223

Character count: 56599

Evaluator Bias in Performance Evaluation of Online Transportation Driver: Empirical Study in Indonesia

35

LUFY YUWANA MURSITA*

Sekolah Tinggi Ilmu Ekonomi Perbanas Surabaya

Abstract: Online transportation ⁴⁷ a developing industry in Indonesia applies bonus-or-terminate incentives based on both objective and subjective performance evaluation. Regarding bias problems often found in the subjective evaluation, this paper aims to examine factors that influence bias in performance evaluation of the online transportation drivers. The data were collected by an online survey to users of online transportation in Indonesia using convenience sampling. Multiple regression analysis was utilized to analyze 163 data. The result shows that users of online transportation services tend to generate a biased rating, which is leniency bias. Altruism and knowledge of incentive ²⁹ es scheme as the users' internal factors significantly affect the biased evaluation. On the other hand, the external factors of the users, i.e., the frequency of usage and evaluation timing, do not significantly affect the biased evaluation. However, the other external factor, i.e., travel distance and duration, is found to affect the leniency bias positively. This paper concludes that more information related to the evaluation object and the inherent characteristic of an individual as the effect of collectivist national culture may lead to the generation of biased performance evaluation by the evaluator to help evaluated party avoiding penalty/termination in the competitive working environment.

Keywords: Evaluation Bias, Online Transportation, Leniency Bias, Altruism

Abstrak: Transportasi daring sebagai industri yang sedang berkembang di Indonesia menerapkan insentif bonus-or-terminate berdasarkan evaluasi kinerja objektif dan subjektif. ³⁴ rkaitan dengan masalah bias yang sering ditemukan dalam evaluasi subjektif, penelitian ini bertujuan untuk menguji faktor-faktor yang menentukan adanya bias dalam evaluasi kinerja pengemudi transportasi daring. Data dikumpulkan melalui survei daring kepada pengguna transportasi daring di Indonesia dengan teknik convenience sampling. Analisis regresi berganda digunakan untuk menganalisis 163 data. Hasilnya menunjukkan bahwa pengguna layanan transportasi daring cenderung menghasilkan penilaian yang bias, yaitu berupa leniency bias. Altruisme dan pengetahuan skema insentif sebagai faktor internal pengguna ditemukan secara negatif memengaruhi leniency bias. Di sisi lain, faktor eksternal pengguna, yaitu frekuensi penggunaan dan waktu evaluasi, ditemukan tidak memengaruhi bias tersebut. Namun, faktor eksternal lainnya, yakni jarak dan durasi perjalanan, ditemukan berpengaruh positif terhadap leniency bias. Hasil penelitian ini memberikan kesimpulan bahwa banyaknya perolehan informasi terkait dengan objek yang dievaluasi dan adanya karakteristik yang melekat pada individu (yang dapat menjadi indikasi budaya

*Corresponding author: lufi.yuwana@perbanas.ac.id

nasional) dapat mengarahkan pada terjadinya evaluasi kinerja yang bias yang sebenarnya dimaksudkan untuk membantu pihak yang dievaluasi agar terhindar dari sanksi/penghentian dalam persaingan kompetitif di lingkungan kerjanya.

Kata Kunci: Altruism; Bias Evaluasi; Leniency Bias; Transportasi Daring

44
1. Introduction

The use of subjectivity is the result of the consideration that objective measures do not adequately represent the actual performance of employees (Gibbs *et al.*, 2004; Merchant and Van der Stede, 2017; Zájbojník, 2014). This type of incentive is best suited for complex work environments, where job design involves multiple tasks and decision-making, as well as an unpredictable environment (Gibbs *et al.*, 2004). The online transportation industry meets the criteria of an unpredictable environment since more companies enter the market, and competitors' innovation becomes more aggressive, so the users' bargaining power tends to get bigger.

Subjectivity is widely used in performance measurement and evaluation, both as a primary and complementary measure, in addition to objective (formulaic) measures. This subjective-weighted performance evaluation is applied to online transportation in Indonesia, such as Go-Jek and Grab. The performance of an online transportation driver is determined by two measures, i.e., formula-based, in the form of multiplied service points in each of the service types, and the subjective rating which is given by customers (Go-Jek, 2018; Grab, 2017). Both of them determine the level of performance and the number of incentives received by the driver. The achievement of formula-based incentives determines the number of daily bonus can be obtained by the drivers. The number of received orders in a day is translated into accumulated service points to be compared with the minimum threshold for obtaining the bonus. If the drivers pass the minimum points, they have an opportunity to get a bonus. However, the opportunity is canceled out if the drivers' performance rating below the prescribed standard level. A poor rating indicates low customer satisfaction and unmet service quality. Therefore, in addition to failure to receive the bonus, the under-performing drivers may also receive a penalty (i.e., account suspension) or even termination. Thus, the online

transportation industry specifically applies subjective evaluation by service users to complete the deficiency in formula-based evaluation, which cannot assure the quality of the service. According to its nature, the incentives scheme in place is called “bonus-or-terminate” (Maestri 2014), which combines both positive and negative incentives in a system.

The use of subjectivity in online transportation performance evaluations and incentives makes “customers are managers” because the rating they provide determines the incentives for the driver (Nastiti, 2018). However, this subjective rating system by customers potentially causes some problems. The poor practice of performance evaluation can lead to the dysfunctional behaviors of the drivers, such as falsifying service orders to obtain a perfect rating (Zaenudin, 2018), which finally impacts on driver’s reluctance to work (Nastiti, 2018). Furthermore, the driver’s final rating comes from evaluations by different evaluators. Nonetheless, the information for evaluators is limited to what they know during the duration of using the driver’s service, which cannot be repeatedly learned by a customer from the same driver in a row. It means that the necessary information to provide the drivers’ performance evaluation is inadequate because it is only obtained from an event at a single point of time (cross-sectional). This condition indicates the inefficiency and tendency to bias. Thus, this also leads to a decrease in drivers’ work motivation.

Previous research has shown that subjective evaluation can lead to bias, such as favoritism (Ittner *et al.*, 2003; Prendergast and Topel, 1993, 1996), centrality bias (Bol 2011; Moers 2005), leniency bias (Moers, 2005; Prendergast and Topel, 1993), outcome effects (Long, Mertins, and Vansant 2015; Mertins, Salvador, and Long 2013; Ghosh 2005), and the hindsight effect (Hawkins and Hastie, 1990), where performance evaluations are perceived to be non-fair and inconsistent for employees. These biases cause the phenomenon of a discrepancy between the actual performance and the given rating, either less or more, even when the total frequency of the evaluation is high. Two systematic biases which have the effect of suppressing and inflating employee performance ratings are sequentially referred to as centrality bias and leniency bias (Bol 2011; Moers 2005). In the end, these biases have the same negative impact on the

sluggishness of the work of online transportation drivers. The more extended use of these biased, subjective weighting triggers the question of what actual goals and impacts of evaluation system implementation are expected for employees and management.

Research has found that leniency bias has a more complicated effect on an agent's performance than the centrality bias (Golman and Bhatia, 2012). Lenient rating from the evaluator leads to a higher expected wage of the employee. This bias provides a higher rating and payment for employees regardless of the level of their actual performance. For below-average performers, this rating causes laziness to exert their work effort because they have been able to receive reasonable payment without working harder. On the other hand, for above-average performers, this rating does not have an impact on future performance (Bol 2011). In the short term, leniency bias can cause a higher level of employee satisfaction. However, in the long-term, it can lead to less incentive to work since the resulted rating assures a good wage for all employees. Therefore, leniency bias is the focus of this study. Besides, there is a limited number of studies on leniency bias (Gong *et al.* 2019).

The subjectivity in performance evaluation for incentive determination appears to be problematic and only effective as well as gives more positive impacts to the agent when the evaluator does have sufficient knowledge and ability about the object of evaluation (drivers' performance). Other factors determine the effectiveness of this subjective measure. However, evaluations made by evaluators who are not part of the company, i.e., the users of the transportation services, lead to higher possible bias in the process and evaluation results. This research is interested in finding what variables can determine the existence of a biased rating, especially lenient rating. Therefore, this study specifically aims to examine the determinants of leniency bias in subjective evaluation by the Indonesian users of online transportation. This paper focuses on the evaluator's perspective, who has the discretion to provide a subjective rating.

This research is important as it indicates the services users' behavior as performance evaluators of the online transportation service who tend to be naive. There are only a small number of studies have been conducted to test the bias in subjective performance evaluation from the evaluator's side. This research is also necessary

because most research on incentives is in the context of positive incentives but rarely on negative incentives. In this study, the test is performed in the context of positive and negative incentives applied at the same time. Both types of incentives which determine the behavioral tendency of the subjective performance evaluation by the online transportation users who consider the fact of drivers' (evaluated party) working environment competitiveness is highly interesting to identify.

Through an online survey, the research shows that one of three factors related to the usage of online transportation services is the determinant of leniency bias, i.e., travel distance and duration. The two remaining factors, i.e., frequency of use and evaluation timing, are not significantly correlated with the bias. This first part of the results indicates the importance of enough information about the drivers for users to provide a more accurate subjective rating. This study also finds that altruism and knowledge of the incentive significantly affect leniency bias. This result indicates that internal factors have a stronger effect on the biased, subjective performance evaluation than the external factors. It also reflects the depiction of collectivist national culture on the users' behavior.

This study contributes empirically by showing the behavioral implications of work incentives, which usually seen from an economic perspective. Besides, it also contributes to the behavioral research literature by showing the determinant of the bias of subjective weighted performance evaluation, which can be developed into a more established theoretical framework. It practically contributes to evaluators and companies by pointing out what matters which encourage bias and injustice in the process and results of employees' performance evaluations by service users. Thus, some steps can be taken to anticipate the occurrence of bias, either by improving the incentive scheme or motivating evaluators to evaluate more objectively.

This remainder of this paper is organized as follows. The next section consists of an explanation of online transportation and hypotheses development based on a theoretical framework. Then the following section is an overview of the method for the study and followed by the results and the discussion. This paper closes with the explanation of conclusion, implication, limitation, and suggestions for future research.

2. Theoretical Framework and Hypothesis Development

2.1 The History and Development of Online Transportation in Indonesia

Online transportation in Indonesia has developed since the idea to empower online *ojek*. *Ojek* or motorbike taxi is riding service provided by motorcyclists to customers who want to reach their ordered destination. It is one means of transportation available in this observed country. There are many *ojek* in Indonesia due to the high number of motorcyclists, especially in the cities. Previously, they have some places to be their post for getting customers, such as in front of traditional markets, malls, offices, and others. However, demand for *ojek* is difficult to increase or even not stable, at least, because of two reasons. First, it provides pick up services to a specific destination, which causes fare to be higher than other public transportations, such as city bus and *angkot* (city transport). Second, based on convenience, some people prefer a car taxi to *ojek* because it protects customers from the hot weather in the city. Thus, *ojek* users are mostly people who have subscribed to and regularly use the services. Therefore, the competition is high and causes a lack of work enthusiasm for the drivers.

The problem then inspired an entrepreneur to initiate business, which empowers the *ojek* by utilizing online devices to assist *ojek* services. This service is provided under Go-Jek corporation. It is explained in their website that Go-Jek was established in 2010 as a motorcycle ride-hailing phone service that has evolved into an on-demand mobile platform and a cutting-edge app to provides extensive range services by now with transportation services still becomes the main business (Go-Jek, 2018). Through the online service, the motorcyclist can easily reach the customer and provide customized orders. Customers are also simplified in looking for drivers to fulfill their pickup needs. The pickup and destination point are also can be determined more accurately by the customers, which ease drivers to find both locations.

Drivers join Go-Jek company as partners, not employees, in which they are given identity and license to provide services by using the app and also have many kinds of work support, such as health and accident cover and financial services and insurance (Go-Jek, 2018). Partners differ from employees as they ideally join Go-Jek not to make it as the primary source of income or activity but as an additional one. Unfortunately,

most of the drivers consider this activity as their primary source of income because they understand this as a better work opportunity than the other kinds of jobs in transportation. This service is now available in 50 cities in Indonesia, with over 400,000 total drivers (Go-Jek, 2018).

Besides Go-Jek, the other platforms also joined the industry, i.e., Grab and Uber. Grab and Uber formerly only provided online car-taxi services. By the increasing needs of online *ojek*, they also expand their services to this area of service. Therefore, they build the online transportation industry in Indonesia, with both of them as the most prominent platforms since Uber was acquired by Grab (Keeton-Olsen, 2018). This industry is phenomenal and very helpful to society. Moreover, it also expands its services not only in providing transportation (motorbike and car), but also facilities for shopping (food, medicine, ticket, and mart), payment (mobile credit top-up and bills), and even lifestyle (massage, cleaning, beauty, and automotive).

As previously explained, the drivers' performance is evaluated by two kinds of performance measures; they are objective and subjective measures. The objective measure is based on the number of orders served by the drivers. In contrast, subjective evaluation is based on the rating provided by the customers. The usage of the objective measure has almost no problems. However, subjective evaluation is somewhat controversial since it is given by the parties who do not come from the internal company. The arbitrary rating given by customers can lead to the sluggishness of work or even resignation. It is because customers have limited information to give an evaluation.

On the other hand, the online transportation industry uses a double incentive system (carrot and stick). Good performance leads to more incentives. Otherwise, bad performance causes a penalty until dismissal. This system gives more pressure to the drivers who deal with highly competitive work dynamics due to the high number of drivers employed by the industry.

2.2 Negative Incentives versus Positive Incentives

Most of the existing incentive studies, particularly in the field of experiments, focused on positive incentives (Lourenço *et al.* 2018). In a bonus-or-terminate scheme

that combines both types of incentives, a positive evaluation is expected to come with monetary compensation. In contrast, negative incentives can lead to dismissal (Maestri 2014). Therefore, employees are partly motivated by the threat of dismissal and partly by the expectation of receiving bonuses (Fuchs 2015; Maestri 2014). However, in this case, the determination of the basis for obtaining such incentives is based on performance evaluations by users of online transportation services that have several limitations.

The difference between negative incentives and positive incentives is important because the prospect theory states that people react differently to gains and losses (Tversky and Kahneman, 1991). However, the negativity bias phenomenon shows that people tend to weigh more substantial on the negative information rather than positive information. Similarly, people will avoid a loss (negative incentives) rather than benefit from a gain (positive incentives). That is, negative incentives have a stronger effect than positive incentives. Thus, in the presence of negative incentives, employees increase their efforts in order to avoid adverse situations more than they do to raise positive incentives (Baumeister *et al.* 2001). Employees change their performance to meet the threshold and avoid termination (Lourenço *et al.* 2018).

On the other hand, the existence of goals serves as a reference point in prospect theory in which the outcome is evaluated and classified as either gain or loss (Heath, Larrick, and Wu 1999). In the negative incentives, goals separate the gain (keeping the job) from the loss domain (penalty/termination). Nonetheless, by achieving a goal, people are in a better position to gain a future positive outcome. Otherwise, people who are below the goal will change their habits in order to meet the prescribing level and avoid penalty/termination. Therefore, goals play a role in prospect theory (Lourenço *et al.* 2018). Bringing back to the context of online transportation, the use of negative and positive incentives at the same time may imply two drivers' behavioral tendencies regarding the achievement of the goals itself.

2.3 The Use and Impact of Subjectivity in Performance Evaluation

Management accounting researchers found one of the determinants of the use of subjective incentives is the extent to which the achievement of incentives target with

objective measures is difficult and leads to significant consequences if not met (Gibbs *et al.* 2004). Subjective incentives allow the evaluator to use additional relevant information that emerged during the measurement and evaluation period. Some researchers have found that subjectively weighted incentives aim to reduce perceived weaknesses in quantitative (formula-based) performance measures (Gibbs *et al.* 2004; Zájbojník 2014). Other researchers explained, in the use of subjective evaluations on incentives scheme in conjunction with formula-based (objective) evaluations, there is an influence on each other among those measures. It was found that the level and controllability of objective measures influenced subjective performance evaluations (Bol and Smith, 2011). The lower and more uncontrollable, the higher the use of subjective measures, and vice versa. In general, it can be concluded that the use of subjective incentives can improve alignment between employee and company interests and reduce employee risk by closing the objective measure weakness.

Since subjective measure equips objective measures, its use in determining incentives has a positive impact. As economists explained, the use of subjectively weighted incentives increases the employees' satisfaction, which further increases productivity and corporate profits (Gibbs *et al.* 2004). Other researchers found that incentives with subjective weights can also encourage employee's knowledge sharing behavior because employees will get more benefits or results by doing that (Cheng and Coyte, 2014). From these studies, it is known that the purpose of using subjectivity in performance evaluation and determination of incentives among others is to improve employee's motivation and performance as well as attitudes and positive ways in order to achieve it.

Some researchers revealed that subjectivity means requiring judgment in action, including in assessing and determining incentives (Moers 2005; Gibbs *et al.* 2004; M. J. Gibbs *et al.* 2009). Evaluators have the discretion to provide assessment and evaluation because there are no clear measurement standards. However, some of these researchers also pointed out that if evaluators are unfair and biased, thus subjectivity in performance evaluations can lead to a substantial risk to employees, particularly on satisfaction and performance (Gibbs *et al.* 2004). In general, evaluations and incentives

of subjective measures will only be effective when evaluators are fair and unbiased in judgment.

Economic researchers proved that subjectivity improves satisfaction and performance only when there is sufficient trust between the evaluator and the evaluated party (Gibbs *et al.* 2004). The results of this study are supported by accounting and business researchers who suggested the results of their surveys in the context of the public sector. Subjectivity in the practice of performance evaluation reduces the clarity of the mission according to employees and their trust in the evaluator, thereby decreasing motivation (Van Rinsum and Verbeeten, 2012). Both studies showed that the existence of trust also determines the effectiveness of subjective evaluation and incentives. The existence of trust implies that there is no bias and injustice in the system practice.

2.4 Altruism

Altruism is a voluntary action to help others in issues related to their work. Altruism is one part of Organizational Citizenship Behavior (OCB) (Posdakoff and MacKenzie, 1994). It is fostered, channeled, or impeded by the socialization experience, which varies across cultures (Draguns, 2013). When their altruism is high, people will not monitor the agent's performance even if the agent keeps their best performance (Giebe and Gürtler, 2012), i.e., biased evaluation. Previous studies have found the effect of altruism. It undermines the threat of dismissal but increases the credibility of bonuses or incentives (Dur and Tichem, 2015). In other words, higher altruism may lead to a higher bonus, while productivity may be lower. However, the effectiveness of the incentives system is altered (Van Rinsum and Verbeeten, 2012).

Previous research has investigated altruism. A behavioral study indicated that leniency bias occurs as the result of an evaluation that is used to determine the employee's pay and the level of altruism of the evaluator (Golman and Bhatia, 2012). The high level of altruism can lead to the limited effectiveness of incentives (Van Rinsum and Verbeeten, 2012). Experimental research found that altruistic behavior is more required and less optional for women rather than for men. Thus, gender differs performance evaluation and reward recommendation, even at the same level of altruism

(Heilman and Chen, 2005). Another research showed that altruistic motive mediates the relationship between helping behavior and reward allocation in performance evaluation (Johnson *et al.* 2002).

2.5 Hypotheses Development

The previous research mentioned three things which affect performance rating accuracy, i.e., the performance of the evaluated party itself, the evaluator's observation of the performance which may lead to observation bias, and the evaluator's memory about the result of the observation (memory bias) (Wherry and Bartlett, 1982). This study focuses on the last two points that are beyond the control of the evaluated party or the employee but rely on the evaluator. In the context of online transportation, it is expected that the evaluator should be fair and unbiased in evaluating the driver.

Economic research has stated that inefficiencies in the bonus-or-terminate incentive scheme decrease when employees are evaluated more frequently (Maestri 2014). Gong *et al.* (2019) supported that supervisors tend to provide greater leniency bias to avoid the social and economic cost of providing accurate performance ratings, especially for a low-performer employee under uncertain conditions. Thus, more frequent use of online transportation services leads users to provide a more lenient rating. Besides, since the meeting between a user and a driver is somehow cross-sectional (non-recurring) in nature, as well as the evaluator is the outsider of the drivers' company, the potential consequence in the evaluation is a spontaneous and relatively naive evaluation, given limited information had by the evaluator for consideration. In other words, more frequent use of the services causes increasingly less attention to the evaluation process, which can produce a biased rating. In contrast, less frequent use of services leads to a lower possibility of bias as the attention of the evaluation process is still high. Thus, the following hypothesis is proposed:

H1: The frequency of use of online transportation services positively affects leniency bias.

Evaluation time chosen by the evaluator relates to memory, whether it is immediate or postponed after the driver's service is completed. The theory of memory states that

the remembering process is closely related to the bias in the observed perceived behavior (Wherry and Bartlett, 1982). It explains the tendency of bias when performance evaluation is not done immediately after the employee completes the performance. As a result, there will be a bias when performance evaluation is postponed because some of the existing information and feelings may have been reduced or lost. The ratings given immediately after the performance observation period will be more accurate than those given after a long period (Wherry and Bartlett, 1982). Therefore, evaluation given right after the service will result in more accurate than postponed evaluation. In other words, biased evaluation more likely to occur when there is a time between the service usage and the evaluation. Thus, the following hypothesis is proposed:

H2: Evaluation timing of online transportation services positively affects leniency bias.

The interaction between the evaluator and the evaluated party can affect the effectiveness of performance evaluation (Goffin and Anderson, 2007). The interaction of the principal and agent in online transportation occurs during the services. It means that travel distance and duration affect the evaluator's observation and its effectiveness. Economic researchers found that in order to produce an unbiased subjective evaluation, evaluators should devote time and effort to gather adequate information regarding employee performance. Lack of information and the closeness between the evaluator and the evaluated party cause leniency bias, i.e., loose in assessing, and centralizing bias, i.e., rating around standard values (Bol 2011). Another bias found by other researchers is favoritism, that is, the evaluator acts on personal preference to the employee by preferring one or more employees to the others (Ittner *et al.*, 2003; Prendergast and Topel, 1993, 1996). This bias makes it difficult to distinguish whether the good rating comes from the biased or unbiased evaluation. In the context of this industry, it is possible for evaluators who travel with a long time and distances to obtain more information regarding the driver's performance. Evaluators will have more chances to assess the driver's performance better. Therefore, the distance and travel

duration of service is expected to decrease the bias in evaluating performance. Thus, the hypothesis is proposed as follows:

H3: Travel distance and duration of online transportation services negatively affect leniency bias.

Altruism is the source of asymmetry in avoiding unfair evaluation, which can lead to leniency bias (Golman and Bhatia, 2012). In a workplace setting, which agent's performance is unverifiable and the usage of double incentives (positive and negative), it was found that altruism increases the credibility of positive incentives. At the same time, it decreases the credibility of negative incentives (Dur and Tichem, 2015). When the evaluator's altruism is high, the evaluator will not monitor the agent's performance even if the agent keeps the best performance (Giebe and Gürtler, 2012). The nature of wanting to help others, which in this case is the driver, encourages the evaluator to provide an evaluation rating that may be better than the driver should have based on his performance. The hypothesis is also driven by the fact that Indonesian people are more collectivist than individualist in the perspective of national culture (Hofstede, 2018). The low score of the individualism dimension of Indonesia means the existence of higher social preference rather than individual preference. People tend to help each other because of the feeling of high interdependence, among others. Therefore, the result of the performance evaluation can be more biased. Thus, the following hypothesis is proposed:

H4: The altruism of online transportation services users positively affects leniency bias.

Specific arrangements that facilitate the increase in bias, such as the knowledge that the rating will have a direct effect on the evaluated party, may decrease the accuracy of the evaluator (Wherry and Bartlett, 1982). In the case of online transportation, a good rating will lead to a bonus or reward for the drivers. In contrast, the poor rating will lead to a penalty or dismissal of the drivers. When the evaluators understand how their evaluation will impact the driver, the evaluator will tend to decrease the accuracy of his evaluation to provide a higher rating. Otherwise, when the evaluators are less

knowledgeable regarding the incentives system, they will give a more objective evaluation or even lower than the expected level. The preference to provide a higher and lower rating as the presence of incentives knowledge is due to the intention of users to minimize the negative effect they may cause to the drivers. Marchegiani *et al.* (2016) supported this prediction by stating that failing to reward a high-performer is more detrimental than failing to rewarding low-performer agents, which motivates evaluators to be more lenient. Thus, the hypothesis is proposed as follows:

H5: Knowledge of the incentive scheme of online transportation services users positively affects leniency bias.

11

Based on the hypotheses above, the theoretical framework of this research as follows.

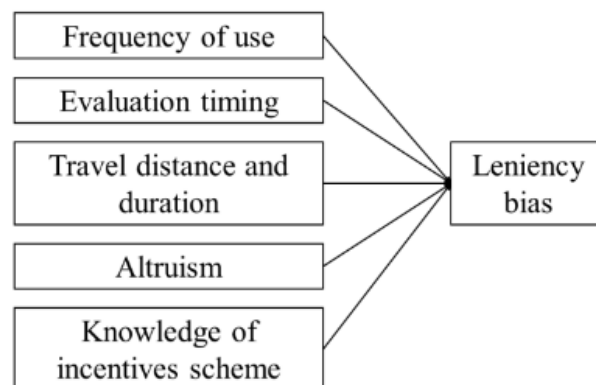


Figure 1.
Theoretical framework

3. Research Method

Data for this study were collected by administering an online survey during May 2018 to reach users of online transportation throughout Indonesia. This study utilized a convenience sampling method by applying criteria that respondents should be between 17-64 years old and have ever used online transportation services. This age criteria refers to the range of productive age in Indonesia (BPS, 2018) and ensures the ability and maturity to provide a performance evaluation. The data were analyzed using multiple regression analysis in SPSS.

This study uses one dependent variable (i.e., performance evaluation bias) and five independent variables (i.e., frequency of service, evaluation timing, travel distance and duration, altruism, and knowledge of incentive scheme). The dependent variable and two independent variables (i.e., altruism and knowledge of incentive scheme) are pre-tested for validity and reliability. The operational definitions and measurements of those variables are explained below.

Performance Evaluation Bias. Performance evaluation bias in this study is the leniency bias, which is defined as a bias that occurs when the rating is given is always high regardless of the quality of the actual performance. The measurement of this variable is based on the total of the result of choice on five statements with a 5-point Likert scale ranging from 1 “strongly disagree” to 5 “strongly agree,” as presented in Appendix” The higher the sum of the five-question items’ score, leniency bias tends more to exist. Cronbach’s alpha of this variable is 0.747 (reliable).

Frequency of Service Use. The frequency of service use is defined as the frequency or number of user’s online transportation usage within a month. The higher the usage, the higher the frequency of use will be. This variable is measured by the respondent’s choice over five categories of service usage amount. The categories include usage frequency range per month: 1-10 times, 11-20 times, 21-30 times, 31-50 times, and more than 50 times.

Evaluation Timing. Evaluation timing is the evaluation given by the user after the expiration of the use of online transportation service. Generally, evaluation time is divided into immediately after service and postponed (with time lag since the completion of the service, usually when the user re-open the application). Evaluation timing is measured by a dummy variable to differentiate the immediate and postponed evaluation. Score 1 is for immediate valuation, while score 0 is for postponed evaluation. This dummy variable is determined based on the respondents’ responses.

Travel Distance and Duration. Travel distance and duration are defined as the length of travel taken by users in using the services of online transportation. The longer the distance kilometers and duration of service use, the higher the travel distance and duration will be. This variable is measured by providing five choices of travel distance

and duration, which are then labeled from 1 to 5 divided based on the number of kilometers and minutes of each travel. Score 1 is for 5-15 minutes (1-5 km) travel, score 2 is for 16-30 minutes (6-10 km) travel, score 3 is for 31-45 minutes (11-15 km) travel, score 4 is for 46-60 minutes (16-20 km) travel, and score 5 is for more than 50 minutes (more than 20 km) travel. The equivalence between travel distance and travel duration uses approximation need for travel in real condition.

Altruism. Altruism is the extent of willingness to help others voluntarily related to their work efforts. This variable is measured by the Smith *et al.* altruism scale (1983) used by (Podsakoff *et al.* 1990), which consists of five statements with a 7-point Likert scale ranging from 1 “strongly disagree” to 7 “strongly agree” as presented in Appendix. Cronbach’s alpha of this variable is 0.755 (reliable).

Knowledge of Incentives Scheme. Knowledge of incentives scheme is the level of understanding of users as evaluators of incentives scheme implemented by online transportation service providers for driver performance. This knowledge is measured by five statements giving the option of a 5-point Likert scale consisting of 1 “do not know at all” to 5 “know for sure” as presented in the Appendix. Cronbach’s alpha of this variable is 0.769 (reliable).

Before hypothesis testing, validity test and reliability test were performed. Subsequently, a descriptive statistical test was performed, and the data were analyzed using multiple linear regression to test the hypothesis. The classical assumption test consisted of tests of normality, heteroscedasticity, and multicollinearity. The level of significance was set at 5%, as suggested in social science research (Sekaran and Bougie, 2016).

4. Results and Discussion

The data were obtained through online surveys filled by users of online transportation in various cities in Indonesia. For further analysis, this research used 163 usable and complete of 172 data from 35 cities and 18 provinces in Indonesia. Table 1 and Table 2 shows the demographic information and descriptive statistics of the

respondents, respectively. The average of the respondents' age is 23.05 years, with the range from 18 to 45 years old. Most users are 21-30 years old. It means that the majority of the users are late adolescence and early adults. About 73.6% of the respondents are women. Most of the respondents graduated from bachelor's degree and senior high school with a percentage of 58.9% and 34.4%, respectively. The majority of the respondents have no working experience (44.8 %) or have working experience of fewer than three years (44.8%). The mean score of the performance evaluation bias variable was 21.23 of 25, indicating the tendency of high leniency bias in performance evaluation.

Table 1.
Demographic Information

Category	No. of respondents	%
Ages		
18-20	30	18.40
21-30	125	76.69
31-40	7	4.29
41-45	1	0.61
Gender		
Man	43	26.38
26 Woman	120	73.62
Education		
Senior high school	56	34.36
Diploma degree	5	3.07
Bachelor degree	96	58.90
18 Master degree	6	3.68
Work Experience		
No experience	73	44.79
< 1 year	36	22.09
1-3 years	37	22.70
3-5 years	8	4.91
> 5 years	9	5.52
Total usable data	163	100

Most respondents have experienced using multiple types of transportation services and platforms. One hundred and thirteen people have used Go-Ride provided by Go-Jek, while 93 people have experienced GrabBike from Grab. In-car service, 107 people have used Go-Car or Go-Taxi provided by Go-Jek, and 105 people have used GrabCar or GrabTaxi from Grab. However, outside of those two large platforms, other platforms

also provide the same services. Thirty-one people have experienced these other platforms' services. This data shows that the existence of various platforms is used interchangeably by the user.

Table 2.
Descriptive Statistics

Variables	Minimum	Maximum	Mean	Std. Dev.
Age	18.00	45.00	23.0491	3.57254
Education	1.00	4.00	2.3190	0.99198
Work experience	0.00	4.00	1.0429	1.17247
Frequency of usage	1.00	5.00	1.3252	0.76887
Travel distance and duration	2.00	8.00	3.7791	1.31480
Altruism	11.00	35.00	26.2515	4.77822
Knowledge of incentives scheme	9.00	25.00	21.8773	3.56210
Leniency bias	15.00	25.00	21.2270	2.18101

Table 3 shows the Pearson correlations matrix for all variables in this research. The correlation coefficient between the frequency of use and leniency bias is 0.162 ($p < 0.05$, two-tailed). Correlations are also found between altruism and knowledge of incentives scheme either altruism and leniency bias with coefficients of -0.012 ($p < 0.01$, two-tailed) and -0.297 ($p < 0.01$, two-tailed), respectively. Knowledge of incentives scheme and leniency bias are correlated at the coefficient of 0.580 ($p < 0.01$, two-tailed). This initial correlation analysis shows that altruism and knowledge of incentives are significantly correlated to leniency bias. Besides, knowledge of incentives scheme is significantly correlated to the altruism. These are the initial findings of this research.

Table 3.
Pearson Correlation Coefficient Matrix

Variables	1	2	3	4	5	6
1. Frequency of use	1.000					
2. Evaluation timing	0.007	1.000				
3. Travel distance and duration	-0.148	0.031	1.000			
4. Altruism	-0.007	-0.087	0.109	1.000		
5. Knowledge of incentives sch.	0.116	-0.010	-0.012	-0.205**	1.000	
6. Leniency bias	0.162*	-0.090	-0.129	-0.297**	0.580**	1.000

** significant at 5%

Before testing the hypotheses, the analysis for non-response bias was performed by comparing the first and last 30 responses (Moore and Tamai, 2002). The result shows

that there is no significant difference between early dan late responses for all variables, except for variable Altruism. However, potential non-response bias for variable Altruism is not a major concern, considering that all other variables are statistically insignificant

The result of the regression analysis is presented in Table 4. The adjusted R square of the regression is 0.382, which means that independent variables explain 38.2% of variances of the dependent variable. The model standard error of the estimate is 1.715 (less than SD=2.181). The F-test is 21.01 ($p=0.000$), which means that the independent variables in the model simultaneously affect Leniency Bias as the dependent variable. Test of control variables was conducted to ensure that factors outside the model do not affect leniency bias. Age, Gender, Education, and Work Experience are found insignificantly correlated with the bias.

Table 4.
Regression Analysis Result

Model	Expected Sign	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		β	Std. Error	Beta		
(Constant)		12.685	1.174		10.801	0.000
Frequency of Use	Positive	0.236	0.178	0.083	1.323	0.188
Evaluation Timing	Positive	-0.311	0.302	-0.064	-1.031	0.304
Travel Distance and Duration	Negative	-0.215**	0.104	-0.129	-2.058	0.041
Altruism	Positive	0.090***	0.029	0.197	3.092	0.002
Knowledge of Incentives Sch.	Positive	0.323***	0.039	0.528	8.304	0.000

** significant at 5%

*** significant at 1%

H1 stated that the Frequency of Online Transportation Usage affects Leniency Bias positively. However, the result shows that the effect is insignificant ($\beta=0.236$, $p=0.188$). This result means that an evaluation bias was not affected by the frequency of online transportation usage. Therefore, H1 was not supported. The result was also not supported H2, which proposed a positive effect of Evaluation Timing on Leniency Bias ($\beta=-0.311$, $p=0.304$). It means that evaluation bias was also not affected by choice of timing in evaluating online transportation driver performance.

H3 stated that Travel Distance and Duration negatively affect Leniency Bias by providing evaluator information to make a judgment about the rating. The result

supports the proposed hypothesis ($\beta=-0.215$, $p<0.041$). This result is consistent with the theory that the longer the travel distance and duration, the more information obtained by the evaluator about the driver's performance. Therefore, the increased information leads to a lower possibility of biased rating.

H4 proposed that Altruism positively affects performance evaluation bias, i.e., Leniency Bias. The result is consistent with the hypothesis ($\beta=0.090$, $p<0.002$). Descriptive results of this variable showed a score of 26.25 of 35.00, which revealed that users tend to be more altruistic, especially towards the driver. It means that altruism is a determinant of leniency bias, where the online transportation users are more willing to help the drivers working in the company. Therefore, H4 was supported.

The result in Table 4 also shows the result of H5, which stated that the Knowledge of Incentives Scheme also has a positive effect on Leniency Bias. The result shows the support for the hypothesis ($\beta=0.323$, $p=0.000$). The descriptive results also revealed that most of the users have known and understood the incentives scheme applied to the driver's working company with a score of 21.88 of 25.00.

Three of five hypotheses in this paper were supported, including the negative effect of travel distance and duration and the positive effect of both altruism and knowledge of incentives scheme on performance evaluation (leniency) bias. The two remaining hypotheses were not supported due to the following possible reasons.

The frequency of use does not significantly affect evaluation bias, meaning that more frequent use using of services does not correlate with the level of evaluation bias. This finding does not support the argument that more frequent usage of services leads to a more lenient rating. Likewise, less frequent usage also does not result in lower leniency bias. Condly *et al.* (2003) suggested that the learning effect may increase performance through the increase in skill and other efforts related to performance. In this context, the intended 'performance' is not the drivers' performance but the users' performance (ability) in providing a subjective evaluation. On the other side, the users who rarely use the services still have great attention to the evaluation process, so they remain trying to provide a more accurate rating. Thus, no significant finding in the relationship between the variables.

Evaluation timing also does not significantly affect the evaluation bias. No difference of leniency bias level between which generated from evaluation with and without time lag after services indicates that users tend to give a consistent rating all the time. It can be the impact of halo theory. Halo effect refers to the tendency to think that person in general as rather good or inferior and make a judgment of the specific attributes by the general feeling (Baker *et al.*, 1994; O'Donnell and Schultze, 2005). Online transportation users generate the next rating based on their perception built on early experience. Then evaluators try to make the current evaluation consistent with the prior performance (Fehrenbacher, Schulz, and Rotaru 2018). That is, when the first or prior experience leads them to provide a good (bad) rating, it can build a perception that online transportation drivers always give excellent (poor) service. Then the users continue to bring the perception to the evaluation of next service orders. If the first impression is good (bad), the subsequent perception can even diminish the bad (good) events effect on the rating because users likely (unlikely) try to understand the underlying reason or situation of the drivers' behavior at the time of service. Thus, evaluation timing does not affect the evaluation bias since the halo effect causes most evaluations to tend to be constantly following the first-time use evaluation. This finding denies the prediction that newer memory will result in more objective evaluation.

The other argument to explain the insignificant effect of evaluation timing is that the limited time users had to provide a rating. Due to the rapid and rush activity, the use of online transportation is intended to help users to expedite their movement. Under this situation, users do not have enough time to determine the exact rating of the driver. Prior literature suggests that intuition mode of subjective evaluation, which is fast and effortless cognitive processing, leads to rapid judgment relied more on prior beliefs and knowledge (Fehrenbacher, Schulz, and Rotaru 2018; Evans 2008). Therefore, evaluation tends to be biased for either given immediately after the services or postponed, or in other words, there is no difference between those two evaluation timings.

More travel distance and duration in H3 provide more opportunity to gather more or even complete information for making a rating decision. In line with Bol (2011),

more accurate transportation comes from judgment based on enough information. More information is obtained through communication during the services, either verbally or nonverbally. This finding is also consistent with Goffin and Anderson (2007). They suggest that more extended interaction among evaluators and evaluated parties results in a more effective evaluation because it provides more opportunities to identify the evaluated party. Therefore, adding in travel distance and duration means a reduction in leniency bias.

As previously defined, altruism is one of personality which refers to a willingness to help other people, regardless of the observed behavior. The support on H4 suggests that more altruistic the users, the more biased evaluation they generate. The evaluation is not genuinely based on the actual performance, but based on the extent of users' empathy, which to help the drivers avoid penalty or termination. Therefore, users choose to give a lenient rating. This result can be potentially supported by the fact of the stable, increasing number of online transportation drivers in Indonesia, which leads to a more competitive working environment for the drivers. Altruistic people were more concerned and paying attention to this fact. They also generate biased evaluations to help the drivers face a harsh working environment. This finding result supports previous research examining altruism effect on behavior. It is consistent with the prior study, for instance, by Golman and Bhatia (2012), which stated that altruism causes leniency bias. In each type of incentive, positive incentives become more credible, but, on the contrary, negative incentives tend to be less credible (Dur and Tichem, 2015). Users do not pay attention any longer to the drivers' effort since the altruistic trait directs users to give a good rating (Giebe and Gürtler, 2012).

Support on H5 reveals that Knowledge of Incentives Scheme positively affects performance evaluation bias, which is leniency bias. Consistent with the prior study (Wherry and Bartlett, 1982), people who know the impact of their evaluation on the evaluated party tend to be more considerate. It results in the high rating assigned for the drivers. Furthermore, this result is also supported by the correlation found between the knowledge of the incentives scheme and altruism performed in the Pearson correlation analysis. Knowledge of the incentives scheme motivates users as the evaluator to give

a more lenient rating for the drivers to save the driver from getting penalty/termination. The knowledge leads to understanding and empathy, which encourages the users to help the drivers avoid negative incentives. This bias generating behavior shows how users respond to the drivers working pressure.

1

5. Conclusion, Implication, and Limitation

The purpose of this study is to examine factors influencing bias in performance evaluation of the online transportation drivers in Indonesia. The results of this study generally support that biased evaluation is more affected by the internal factors of the evaluator, including altruistic personality and knowledge, that encourage them to give higher ratings in general. On the other hand, the external factors of the users, i.e., frequency of use and evaluation timing, do not affect their judgment regarding driver performance evaluation. Travel distance and duration are the only external factor that negatively affects the leniency bias. These results are consistent with most of the previous studies of evaluation bias.

The result of this study is in line with the Indonesian people's characteristics, who are more collectivist and showing helpfulness to each other. Due to the sensitive incentives scheme applied by online transportation operators, the users strive to help drivers to maintain their work by giving above-average performance ratings. Nevertheless, the long-term effect of this users' behavior needs further research to identify whether it leads to the negative impact like have been investigated in the prior study (e.g., Golman and Bhatia, 2012); Bol, 2011).

This study provides either theoretical and practical implications. Theoretically, this research implies that the evaluator's helpful characteristics and deeper information related to the evaluated party can lead to bias in performance evaluation. Meanwhile, in practical terms, this research implies that online transportation providers should be more conservative in interpreting a high rating of the drivers since the presence of leniency bias is identified in the performance evaluation. It means that the company

should provide an education to the users regarding the need to be objective in performing the drivers' evaluation.

This study has several limitations. First, nonresponse bias was found on one independent variable even though this bias was not a major concern. Future research may more consider this bias. Second, this research did not take other kinds of bias (e.g., centrality bias) into consideration, which can potentially affect the provision of performance evaluation. Future research can develop the investigation into the multiple types of evaluation bias. Furthermore, future research can also involve more determinants of performance evaluation bias and build a categorization of those determinants other than internal-external factors.

Reference

- Badan Pusat Statistik. (2018). *Badan Pusat Statistik*. Retrieved April 20, 2018. https://www.bps.go.id/istilah/index.html?Istilah_page=4.
- Baker, George, Robert Gibbons, and Kevin J. Murphy. 1994. "Subjective Performance Measures in Optimal Incentive." *The Quarterly Journal of Economics* 109 (4): 1125–56. <https://doi.org/10.2307/2118358>.
- Baumeister, Roy F., Ellen Bratslavsky, Catrin Finkenauer, and Kathleen D. Vohs. 2001. "Bad Is Stronger Than Good." *Review of General Psychology* 5 (4): 323–70. <https://doi.org/10.1037/1089-2680.5.4.323>.
- Bol, Jasmijn C. 2011. "The Determinants and Performance Effects of Managers' Performance Evaluation Biases." *Accounting Review* 86 (5): 1549–75. <https://doi.org/10.2308/accr-10099>.
- Bol, Jasmijn C., and Steven D. Smith. 2011. "Spillover Effects in Subjective Performance Evaluation: Bias and the Asymmetric Influence of Controllability." *Accounting Review* 86 (4): 1213–30. <https://doi.org/10.2308/accr-10038>.
- Cheng, Mandy M, and Rodney Coyte. 2014. "The Effects of Incentive Subjectivity and Strategy Communication on Knowledge-Sharing and Extra-Role Behaviours." *Management Accounting Research* 25 (2): 119–30. <https://doi.org/10.1016/j.mar.2013.07.003>.
- Dur, Robert, and Jan Tichem. 2015. "Altruism and Relational Incentives in the Workplace." *Journal of Economics and Management Strategy* 24 (3): 485–500. <https://doi.org/10.1111/jems.12099>.
- Evans, Jonathan St. B. T. 2008. "Dual-Processing Accounts of Reasoning, Judgment, and Social Cognition." *Annual Review of Psychology* 59 (1): 255–78. <https://doi.org/10.1146/annurev.psych.59.103006.093629>.
- Fehrenbacher, Dennis D., Axel K.D. Schulz, and Kristian Rotaru. 2018. "The Moderating Role

- of Decision Mode in Subjective Performance Evaluation.” *Management Accounting Research* 41 (March 2016): 1–10. <https://doi.org/10.1016/j.mar.2018.03.001>.
- Fuchs, William. 2015. “Subjective Evaluations: Discretionary Bonuses and Feedback Credibility.” *American Economic Journal: Microeconomics* 7 (1): 99–108. <https://doi.org/10.1257/mic.20130250>.
- Ghosh, Dipankar. 2005. “Alternative Measures of Managers’ Performance, Controllability, and the Outcome Effect.” *Behavioral Research in Accounting* 17: 55–70. <https://doi.org/10.2308/bria.2005.17.1.55>.
- Gibbs, Michael J., Kenneth A. Merchant, Wim A. Van Der Stede, and Mark E. Vargus. 2009. “Performance Measure Properties and Incentive System Design.” *Industrial Relations* 48 (2): 237–64. <https://doi.org/10.1111/j.1468-232X.2009.00556.x>.
- Gibbs, Michael, Kenneth A. Merchant, Wim A. Van Der Stede, and Mark E. Vargus. 2004. “Determinants and Effects of Subjectivity in Incentives.” *Accounting Review* 79 (2): 409–36. <https://doi.org/10.2308/accr.2004.79.2.409>.
- Giebe, Thomas, and Oliver Gürtler. 2012. “Optimal Contracts for Lenient Supervisors.” *Journal of Economic Behavior and Organization* 81 (2): 403–20. <https://doi.org/10.1016/j.jebo.2011.11.003>.
- Goffin, Richard D., and David W. Anderson. 2007. “The Self-Rater’s Personality and Self-Other Disagreement in Multi-Source Performance Ratings: Is Disagreement Healthy?” *Journal of Managerial Psychology* 22 (3): 271–89. <https://doi.org/10.1108/02683940710733098>.
- Goldman, Russell, and Sudeep Bhatia. 2012. “Performance Evaluation Inflation and Compression.” *Accounting, Organizations, and Society* 37 (8): 534–43. <https://doi.org/10.1016/j.aos.2012.09.001>.
- Gong, Na, Wai Fong Boh, Anne Wu, and Tsuilin Kuo. 2019. “Leniency Bias in Subjective Performance Evaluation: Contextual Uncertainty and Prior Employee Performance.” *Emerging Markets Finance and Trade*, 1–15. <https://doi.org/10.1080/1540496X.2019.1660161>.
- Go-Jek. 2018. “Perhitungan Poin dan Bonus Driver GO-JEK Luar Jabodetabek.” Go-Jek, Maret 22. Retrieved April 1, 2018. <https://driver.go-jek.com/hc/id/articles/235509348-Perhitungan-Poin-dan-Bonus-Driver-GO-JEK-Luar-Jabodetabek>.
- Go-Jek. 2018. “About Go-Jek.” Retrieved August 1, 2018. <https://www.go-jek.com/about/>
- Grab Indonesia. 2017. “Grab – Cara Kerja Bintang/Rating.” Artikel Grab Indonesia, Juli 13. Retrieved April 1, 2018. <https://www.grab.com/id/blog/driver/grab-cara-kerja-bintang-rating/>.
- Hawkins, Scott A, and Reid Hastie. 1990. “Hindsight: Biased Judgments of Past Events After the Outcomes Are Known.” *Psychological Bulletin* 107 (3): 311–27.
- Heath, Chip, Richard P. Larrick, and George Wu. 1999. “Goals as Reference Points.” *Cognitive Psychology* 38 (1): 79–109. <https://doi.org/10.1006/cogp.1998.0708>.
- Heilman, Madeline E., and Julie J. Chen. 2005. “Same Behavior, Different Consequences: Reactions to Men’s and Women’s Altruistic Citizenship Behavior.” *Journal of Applied Psychology* 90 (3): 431–41. <https://doi.org/10.1037/0021-9010.90.3.431>.

- Hofstede, G. 2018. "What about Indonesia?". Hofstede Insight Website. Retrieved September 12, 2018. <https://www.hofstede-insights.com/country-comparison/indonesia/>.
- Ittner, Christopher D., David F. Larcker, and Marshall W. Meyer. 2003. "Subjectivity and the Weighting of Performance Measures: Evidence from a Balanced Scorecard." *Accounting Review* 78 (3): 725–58. <https://doi.org/10.2308/accr.2003.78.3.725>.
- Johnson, Diane E., Amir Erez, D. Scott Kiker, and Stephan J. Motowidlo. 2002. "Liking and Attributions of Motives as Mediators of the Relationships between Individuals' Reputations, Helpful Behaviors, and Raters' Reward Decisions." *Journal of Applied Psychology* 87 (4): 808–15. <https://doi.org/10.1037/0021-9010.87.4.808>.
- Keeton-Olsen, D. 2018. "Grab officially takes control of Uber's Southeast Asia operations." *Forbes*, March 26. Retrieved August 1, 2018. <https://www.forbes.com/sites/daniellekeetonolsen/2018/03/26/grab-officially-takes-control-of-ubers-southeast-asia-operations/#6dea51796c57>
- Long, James H, Lasse Mertins, and Brian Vansant. 2015. "The Effects of Firm-Provided Measure Weightings on Evaluators' Incorporation of Non-Contractible Information" 27 (1): 47–62. <https://doi.org/10.2308/jmar-50837>.
- Lourenço, Sofia M., Jeffrey O. Greenberg, Melissa Littlefield, David W. Bates, and V. G. Narayanan. 2018. "The Performance Effect of Feedback in a Context of Negative Incentives: Evidence from a Field Experiment." *Management Accounting Research* 40 (January): 1–14. <https://doi.org/10.1016/j.mar.2018.01.002>.
- Maestri, Lucas. 2014. "The Efficiency of Bonus-or-Terminate Incentive Schemes under Subjective Evaluations." *Games and Economic Behavior* 87: 412–18. <https://doi.org/10.1016/j.geb.2014.05.015>.
- Marchegiani, Lucia, Tommaso Reggiani, and Matteo Rizzolli. 2016. "Loss Averse Agents and Lenient Supervisors in Performance Appraisal." *Journal of Economic Behavior and Organization* 131: 183–97. <https://doi.org/10.1016/j.jebo.2016.07.022>.
- Merchant, Kenneth A, and Wim A Van der Stede. 2017. *Management Control Systems: Performance Measurement, Evaluation, and Incentives*. 4th ed. Harlow: Pearson Education Limited.
- Mertins, Lasse, Debra Salvador, and James H Long. 2013. "The Outcome Effect – A Review and Implications for Future Research." *Integrative Medicine Research* 31 (1): 2–30. <https://doi.org/10.1016/j.acclit.2013.06.002>.
- Moers, Frank. 2005. "Discretion and Bias in Performance Evaluation: The Impact of Diversity and Subjectivity" 30: 67–80. <https://doi.org/10.1016/j.aos.2003.11.001>.
- Nastiti, Aulia. (2018). "Saat Gojek dan Grab Berubah Jadi Ajang Eksploitasi Pengemudi." *Pikiran Rakyat*, 28 Maret. Retrieved April 1, 2018. <http://www.pikiran-rakyat.com/nasional/2018/03/28/saat-gojek-dan-grab-berubah-jadi-ajang-eksploitasi-pengemudi-422033>.
- O'Donnell, Ed, and Joseph J Schultz. 2005. "The Halo Effect in Business Risk Audits: Can Strategic Risk Assessment Bias Auditor Judgment about Accounting Details?" *The Accounting Review* 80 (3): 921–39. <https://doi.org/10.2307/4093182>.

- Podsakoff, Philip M., Scott B. MacKenzie, Robert H. Moorman, and Richard Fetter. 1990. "Relationship among Leadership, Organizational Commitment, and OCB in Uruguayan." *Leadership and Organization Development Journal* 1 (2): 107–42. [https://doi.org/10.1016/1048-9843\(90\)90009-7](https://doi.org/10.1016/1048-9843(90)90009-7).
- Podsakoff, Philip M., and Scott B. MacKenzie. 1994. "Organizational Citizenship Behaviors and Sales Unit Effectiveness." *Journal of Marketing Research* 31 (3): 351. <https://doi.org/10.2307/3152222>.
- Prendergast, Canice, and Robert Topel. 1993. "Discretion and Bias in Performance Evaluation." *European Economic Review* 37: 355–65. [https://doi.org/10.1016/0014-2921\(93\)90024-5](https://doi.org/10.1016/0014-2921(93)90024-5).
- Prendergast, Canice, and Robert H Topel. 1996. "Favoritism in Organizations." *Journal of Political Economy* 104 (5): 958–78. <https://doi.org/10.1086/262048>.
- Rinsum, Marcel Van, and Frank H.M. Verbeeten. 2012. "The Impact of Subjectivity in Performance Evaluation Practices on Public Sector Managers Motivation." *Accounting and Business Research* 42 (4): 377–96. <https://doi.org/10.1080/00014788.2012.653747>.
- Sekaran, Uma, and Roger Bougie. 2016. *Research Methods for Business*. 7th ed. West Sussex: John Wiley & Sons Ltd. <https://doi.org/10.1017/CBO9781107415324.004>.
- Tversky, Amos, and Daniel Kahneman. 1991. "Loss Aversion in Riskless Choice: A Reference-Dependent Model." *Quarterly Journal of Economics* 106 (4): 1039–61. <https://doi.org/10.2307/2937956>.
- Van Rinsum, M., and Verbeeten, F. H. M. 2012. The impact of subjectivity in performance evaluation practices on public sector managers motivation. *Accounting and Business Research*, 42(4), 377–396. <https://doi.org/10.1080/00014788.2012.653747>
- Wherry, Robert J., and C. J. Bartlett. 1982. "The Control of Bias in Ratings: A Theory of Rating." *Personnel Psychology* 35 (3): 521–51. <https://doi.org/10.1111/j.1744-6570.1982.tb02208.x>.
- Zábojník, Ján. 2014. "Subjective Evaluations with Performance Feedback." *RAND Journal of Economics* 45 (2): 341–69. <https://doi.org/10.1111/1756-2171.12054>.
- Zaenudin, Ahmad. 2018. "Nasib Driver Ojek Online Tergantung "Bintang"." *Tirto.id*, February 2. Retrieved June 20, 2018. <https://tirto.id/nasib-driver-ojek-online-tergantung-bintang-cEaw>.

Appendix

Altruism

Question	1	2	3	4	5	6	7
1. In the workplace/school, I help others who have been absent.							
2. In the workplace/school, I help others who have heavy workloads.							
3. In the workplace/school, I help orient new people even though it is not required.							
4. In the workplace/school, I will help others who have work-related problem.							
5. In the workplace/school, I am always ready to lend a helping hand to those around me.							

Description: 1 = strongly disagree, 7 = strongly agree

Knowledge of incentives theme

Question	1	2	3	4	5
1. I know that the income of online transportation drivers is based on the number of passengers or the services they obtain					
2. I know that online transportation driver incentives/bonuses are based on the number of passengers or more services obtained after the minimum target is reached.					
3. I know that the service rating that I provide as a passenger greatly influences driver income.					
4. I know that a bad rating (for example, 1 star) can cause a driver to be suspended or unable to provide service for a certain period of time.					
5. I know that a bad rating given by several passengers can cause drivers to be terminated by the company.					

Description: 1 = do not know at all, 5 = know for sure

Leniency bias

Question	1	2	3	4	5
1. I always give a high rating regardless of the level of the actual performance of the driver.					
2. I always give a high rating even though there is an unpleasant attitude or behavior of the driver.					
3. I always give a high rating even though the driver picks up a bit longer than estimated in the application.					
4. I always give a high rating even though there are SOPs (standard procedures) left by drivers, for example, not offering masks.					
5. I always give a high rating even though the driver only speaks as needed to me.					

Description: 1 = strongly disagree, 5 = strongly agree

Cek Plagiarisme Publikasian

ORIGINALITY REPORT

9%

SIMILARITY INDEX

7%

INTERNET SOURCES

6%

PUBLICATIONS

2%

STUDENT PAPERS

PRIMARY SOURCES

1	ijar-iaikapd.or.id Internet Source	2%
2	Sofia M. Lourenço, Jeffrey O. Greenberg, Melissa Littlefield, David W. Bates, V.G. Narayanan. "The performance effect of feedback in a context of negative incentives: Evidence from a field experiment", Management Accounting Research, 2018 Publication	1%
3	www.tandfonline.com Internet Source	<1%
4	www.pfingsttagung.unizh.ch Internet Source	<1%
5	Submitted to iGroup Student Paper	<1%
6	hdl.handle.net Internet Source	<1%
7	Amira Abd-El Aziz Mohamed Riad, Eglal Hafez, Eahab Elsaid, Nicole Taylor. "The Effect of Employee Engagement on Organizational Citizenship Behaviour: An	<1%

Empirical Study on an Egyptian University", Business and Management Research, 2019

Publication

8	pinnacle.allenpress.com Internet Source	<1 %
9	www.econstor.eu Internet Source	<1 %
10	Submitted to Widener University Student Paper	<1 %
11	www.unlv.edu.sg Internet Source	<1 %
12	Submitted to Erasmus University of Rotterdam Student Paper	<1 %
13	www.acrwebsite.org Internet Source	<1 %
14	Carmen Aranda, Javier Arellano, Antonio Davila. "Subjective bonuses and target setting in budget-based incentive contracts", Management Accounting Research, 2019 Publication	<1 %
15	Submitted to Georgetown University Student Paper	<1 %
16	www.baspcan.org.uk Internet Source	<1 %
17	Submitted to University of Leicester Student Paper	<1 %

18	www.cdri.org.kh Internet Source	<1 %
19	pt.scribd.com Internet Source	<1 %
20	dro.deakin.edu.au Internet Source	<1 %
21	res.mdpi.com Internet Source	<1 %
22	Golman, Russell, and Sudeep Bhatia. "Performance evaluation inflation and compression", <i>Accounting Organizations and Society</i> , 2012. Publication	<1 %
23	Gibbs, Michael, Kenneth A. Merchant, Wim A. Van der Stede, and Mark E. Vargus. "Determinants and Effects of Subjectivity in Incentives", <i>The Accounting Review</i> , 2004. Publication	<1 %
24	bura.brunel.ac.uk Internet Source	<1 %
25	link.springer.com Internet Source	<1 %
26	docplayer.net Internet Source	<1 %
27	lumsa.it Internet Source	<1 %

28	pdfs.semanticscholar.org Internet Source	<1 %
29	archiv.ub.uni-heidelberg.de Internet Source	<1 %
30	Kenneth A. Merchant, Wim A. Van der Stede, Thomas W. Lin, Zengbiao Yu. "Performance Measurement and Incentive Compensation: An Empirical Analysis and Comparison of Chinese and Western Firms' Practices", <i>European Accounting Review</i> , 2011 Publication	<1 %
31	id.scribd.com Internet Source	<1 %
32	etheses.dur.ac.uk Internet Source	<1 %
33	lirias.kuleuven.be Internet Source	<1 %
34	www.scribd.com Internet Source	<1 %
35	sinta3.ristekdikti.go.id Internet Source	<1 %
36	Rebekka L. Erks, Joseph A. Allen, Lynn K. Harland, Kelly Prange. "Do Volunteers Volunteer to do More at Work? The Relationship Between Volunteering, Engagement, and OCBs", <i>VOLUNTAS: International Journal of Voluntary and</i>	<1 %

Nonprofit Organizations, 2020

Publication

37

Mary Dana Laird, James J. Zboja, Gerald R. Ferris. "Partial mediation of the political skill-reputation relationship", *Career Development International*, 2012

Publication

<1%

38

www7511.ssldomain.com

Internet Source

<1%

39

aaajournals.org

Internet Source

<1%

40

onlinelibrary.wiley.com

Internet Source

<1%

41

Hougyun Kim. "Transformational Leadership and Organisational Citizenship Behavior in the Public Sector in South Korea: The Mediating Role of Affective Commitment", *Local Government Studies*, 2012

Publication

<1%

42

theses.gla.ac.uk

Internet Source

<1%

43

orca.cf.ac.uk

Internet Source

<1%

44

Dennis D. Fehrenbacher, Axel K.-D. Schulz, Kristian Rotaru. "The moderating role of decision mode in subjective performance evaluation", *Management Accounting*

<1%

Research, 2018

Publication

45

Dur, Robert, and Jan Tichem. "Altruism and Relational Incentives in the Workplace : Altruism in the Workplace", Journal of Economics & Management Strategy, 2015.

Publication

<1%

46

"Management Control and Uncertainty", Springer Science and Business Media LLC, 2014

Publication

<1%

47

Narisa Tianjing Dai, Xi (Jason) Kuang, Guliang Tang. "Differential Weighting of Objective Versus Subjective Measures in Performance Evaluation: Experimental Evidence", European Accounting Review, 2016

Publication

<1%

Exclude quotes Off

Exclude matches Off

Exclude bibliography On