LPEM-FEBUI Working Paper - 014
November 2017

ISSN 2356-4008

WHEN OFFICIALS DON'T KNOW WHAT THEY DON'T KNOW: DUNNING-KRUGER EFFECT IN THE CASE OF GREEN BUDGETING FOR LOCAL GOVERNMENT

Alvin Ulido Lumbanraja

LPEM-FEBUI Working Paper 014

Chief Editor : Riatu M. Qibthiyyah
Editors : Kiki Verico
Setting : Rini Budiastuti

Institute for Economic and Social Research
Faculty of Economics and Business
Universitas Indonesia (LPEM-FEB UI)

Salemba Raya 4, Salemba UI Campus Jakarta, Indonesia 10430
Phone : +62-21-3143177
Fax : +62-21-31934310
Email : lpem@lpem-feui.org
Web : www.lpem.org

LPEM-FEB UI Working Paper 014, November 2017
ISSN 2356-4008

When Officials Don't Know What They Don't Know: Dunning-Kruger Effect in the Case of Green Budgeting for Local Government

Alvin Ulido Lumbanraja [1]*

Abstract
This paper extends the key findings of Kruger & Dunning (1999), which shows that people who are unskilled in a given domain tend to be unaware of their lack of skills, to government circle that is supposed to be filled by professionals. This paper compared individual government officials' self-assessment of their offices' ability to perform certain tasks related to green budgeting with their responses to questions that implicitly assess their actual ability to perform such tasks. Consistent with Kruger & Dunning (1999), individuals who have sufficient knowledge and expertise in a given domain tend to have more accurate self-assessment when asked to rate their own expertise, and vice versa. This paper also discusses the theoretical underpinning of how compensation structure is related with Dunning-Kruger effect on policy design and how tying the outcome with compensation can promote learning and better metacognitive abilities, even for less knowledgeable individuals..

JEL Classification: D86; H10; J30; J45; M52

Keywords
Dunning-Kruger Effect — Green Budgeting — Government Officials

[1] *Institute for Economic and Social Research, Faculty of Economics and Business, Universitas Indonesia (LPEM- FEB UI)*
***Corresponding author**: Campus UI Salemba, Salemba Raya St., No. 4, Jakarta, 10430, Indonesia. Email: alvin.ulido@gmail.com

"The fool doth think he is wise, but the wise man knows himself to be a fool"
- Shakespeare

1. INTRODUCTION

Are *all* government officials, who often deal with uncertainties and ambiguities, better in knowing the limits of their own understanding than average people? Or do they, just like people they serve, tend to overestimate their own level of knowledge when they actually do not know as much?

The propensity to overestimate level of knowledge in the field that one does not actually master or to more accurately (and sometimes underestimate) their level of knowledge in the field that one actually master is known as Dunning-Kruger effect in the field of psychology. While this specific cognitive bias has only been scientifically scrutinized since the introduction of seminal paper of Kruger & Dunning (1999), philosophers throughout centuries have recognized that poor performers tend to overestimate their ability, and vice versa. Bertrand Russell, quoted from Ehrlinger et al. (2008), succinctly and aptly summarized this when, in *New Hopes for a Changing World*, he remarked "One of the painful things about our time is that those who feel certainty are stupid, and those with any imagination and understanding are filled with doubt and indecision".

Although explanations and empirical tests on Dunning-Kruger effect are well-established, most of the existing literatures tend to focus on people who are non-professionals, i.e. non-experts whose livings do not predicate on mastering knowledge in a given field. This might be driven by assumption that professionals have to gain sufficient knowledge in a given field in order to obtain their job in the first place and get paid to perform their duties. One of experiment from Kruger & Dunning (1999), which shows that logical training can improve metacognitive skills, can be interpreted as (supposedly) positive relationship between education and relevant work experience are also supposed to make people more aware of their own level of knowledge.

Indeed, concern about meta-ignorance in professional setting has already been aired in other fields, although government circle has never been explicitly researched before. In the field of academic research, where professional researchers tend to self-regulate, Regehr & Eva (2006) and Huang (2013) noted that reliance on self-assessment among researchers to assess the limits of their own knowledge and skills, a group that is supposed to be experts, may need to be reconsidered. Both argued that even experts as humans can still suffer from cognitive biases, including Dunning-Kruger effect.

Eva & Regehr (2005) and Regehr & Eva (2006) noted that while the increased pace of new medical researches and knowledge requires medical professionals to better assess the area of knowledge that they should learn more, many medical professionals' self-assessment correlate poorly with their actual knowledge. Huang (2013) argued that the existence of Dunning-Kruger effect among academic researchers and peer-reviewers create some sort of echo chamber, in which they suppress their own ignorance by relying on their existing knowledge and rejecting alternative or competing hypotheses in face of new evidences, and ultimately stifles diversity of ideas. This, in turn, strengthen the suggestion that professionals, rather than being group of people who are all capable of self-criticism, can be just as prone to

meta-ignorance.

This suspicion about the existence of meta-ignorance among less knowledgeable individuals in government circle was further prompted by observations on civil servants at subnational level in Indonesia. As part of LPEM's grant from Millennium Challenge Account Indonesia (part of Millennium Challenge Corporation), we have conducted various efforts to increase capacity, including but not limited to trainings and technical assistances, on green budgeting for local government, both at province and municipal (city/regency) levels for almost two years. Before we commenced our training sessions, we found that many province-level and/or regency-level officials said that they are aware of climate change issues and green budgeting, and put blame on limited funds for lack of actions in mitigating climate change to (i) absence of regulatory mandate from Ministry of Interior, (ii) lack of willingness from elected Governors/Mayors/Regents and Regional House of Representatives, and (iii) lack of fund due to other more urgent priorities.

However, when training sessions commenced, it became apparent that those same officials did not actually possess enough relevant knowledge to understand and implement green budgeting, such as ability to reasonably assign monetary value to non-monetary goods like CO_2 emission reduction. This disconnect between officials' perceived ability and their actual ability is puzzling, particularly as local governments have been mandated by law to prepare Regional Action Plan for Greenhouse Gas Emission Reduction (RAD-GRK) since 2011. Given that all aspects of local government are affected by RAD-GRK, local civil servants should have learned enough to accurately rate their own knowledge about the subject matter. Furthermore, as one of experiment in Kruger & Dunning (1999) shows that logical training improves metacognitive skills, and that senior civil servants tend to be better educated and far more experienced than general population, those officials are supposed to be even less prone to Dunning-Kruger effect.

This line of argument, as observer may note, rests on two underlying implicit assumptions. First, senior civil servants that are tasked to deal with green budgeting are supposed to, at least partially, be appointed or promoted to their respective post by merit. In this case, we assumed that senior civil servants have sufficient knowledge/skills that is/are required of his/her appointed post. This is not unreasonable, given that civil servants need to pass relevant civil service entrance examination to join, for example, Office of Environmental Affairs (*Badan Lingkungan Hidup Daerah*) or Office of Transport (*Dinas Perhubungan*). Both the civil servants that are appointed and person who promote said civil servants to their respective posts are therefore assumed to have adequate metacognitive skills; they are assumed to both know that the promoted person is capable or can be easily learn on-the-job to acquire capability in due course.

This study is therefore designed as a preliminary research and stepping stone for further studies on meta-ignorance in government circle, particularly when no clearly defined, quantitative measurement for knowledge evaluation is readily available. To that end, the first part of this paper is devoted to discussion of existing literatures on Dunning-Kruger effect and metacognition. The next part of this paper discusses the methodology, key findings of our research, and limitation of this study. The last section discusses the policy implication of meta-ignorance for public service delivery in general and for green budgeting in particular. Specifically, this paper is interested in what, if any, incentives can be implemented to encourage government officials to improve their metacognitive skills.

2. THEORETICAL FRAMEWORK AND LITERATURE REVIEW

2.1 Concept of Metacognition and Dunning-Kruger Effect

To understand the cognitive mechanism behind why less competent individuals tend to overestimate their level of knowledge and/or competence, we need to first understand the concept of metacognition. The term metacognition, defined very simply, is cognition about cognition or a second-order cognition; thoughts about thoughts, knowledge about knowledge, or reflection about actions (Papaleontiou-Louca, 2008:2). A basic, non-technical, example about metacognition is one's knowledge or awareness about one's level of knowledge in a given field (e.g. "I am aware that I know very little about the concept of metacognition", or "I do not know that I only know so little about psychology").

John H. Flavell, who coined the term metacognition, divided it into two parts, metacognitive knowledge and metacognitive experience. Flavell (1979) defined metacognitive knowledge as "one's stored knowledge or beliefs about oneself and others as cognitive agents, about tasks, about actions or strategies, and about how all these interact to affect the outcomes of any sort of intellectual enterprise", and metacognitive experience/regulation as "conscious cognitive or affective experiences that occur during the enterprise and concern any aspect of it—often, how well it is going".

Lai (2011), further refined the constituent elements of metacognition that were proposed by Flavell (1979) by incorporating insights researches that follows Flavell (1979). Lai (2011) categorized the components of metacognition into cognitive knowledge and cognitive regulation; he summarizes cognitive knowledge, broadly defined, as knowledge about one's own knowledge (Paris & Winograd, 1990; Schraw et al., 2006), others' (Flavell, 1979) and about epistemological understanding in general (Kuhn & Dean, 2004)[1]. Cognitive regulation is also broadly defined as perception or experience of cognition that serves as "quality control" (Flavell, 1979) and process of planning, monitoring, and evaluating cognition (Cross & Paris, 1988; Paris & Winograd, 1990; Schraw et al., 2006; Schraw & Moshman, 1995; Whitebread et al., 2009)[2].

It may become apparent by now that ability to know the limit of one's own knowledge (cognitive knowledge) and ability to create strategy to widen one's horizon is important to gain new knowledge. Researchers has linked metacognition to intelligence; Sternberg (1984) even considered

[1] Cognitive knowledge can be expressed by the ability to accurately respond to question such as "Do I know this?"

[2] Cognitive regulation can be defined by how individual monitor, assess and make strategy to improve their knowledge. This can be expressed by questions such as "How am I doing now? What should I do to improve my skills or knowledge?"

that metacognition, which enable someone to appropriately allocate cognitive resources for learning, as central to intelligence. Within the context of knowledge acquisition, we can infer that when someone has vast knowledge about certain field, he or she tend to know better and more accurately about the extent of available knowledge and limit of his current knowledge, what concepts he should learn to gain more knowledge, and how to do so. In some cases, people with vast knowledge may even underestimate their own knowledge and capabilities, fallen prey to false-consensus effect (Ross et al., 1977)[3].

Conversely, without sufficient knowledge about related field and sufficient humility to admit ignorance, people may tend to overestimate their knowledge on a given field. We can demonstrate the presence of this phenomenon by asking non-economist about their knowledge on "efficient market". To people with no background in financial economics, the phrase "efficient market" may evoke false familiarity, given the seemingly non-technical appearance of the phrase. This may prompt the interlocutor to think of spurious definitions of efficient market, such as 'a market that utilize resources with least amount of waste'. On the other hand, people with some comprehension of financial economics will correctly point out features of efficient market, such as 'securities price fully reflect all available information at that time', but may hesitate in offering comprehensive definition due to awareness that they may not fully know the details of Efficient Market Hypothesis, such as the difference between weak EMH and strong EMH.

This very phenomenon is what constitutes Dunning-Kruger effect; people with low level of knowledge tend to overestimate their ability in that field precisely because they are not actually familiar with the field in question. People who do not know what they do not know tend to think that they almost know all there is to know. Kruger & Dunning (1999) essentially argued that lack of knowledge produces some sort of vicious cycle; one cannot measure accurately his/her level of knowledge without accumulating sufficient knowledge, but the very lack of knowledge hinders people from realizing that they know so little.

The good news is that, as aforementioned, deficit in metacognitive skills (and metacognitive skills in general) is not a permanent condition; training and feedbacks can improve metacognitive skills, i.e. better ability to know the extent and limit of one's own knowledge or competence.

However, as pointed out by Kruger & Dunning (1999), if trainings and feedbacks improve metacognitive skills, and that life experience should have demonstrated that they are ignorant and/or unskilled, why do people with less knowledge or competence still fail to realize their shortcomings? Kruger & Dunning (1999) suggests that (i) people who are not competent seldom receive feedbacks, (ii) some tasks and/or settings preclude people from receiving feedbacks of their shortcomings and why they are not performing well, (iii) even if people receive feedbacks that point to their lack of knowledge or skills, they may attribute it to other factors (Brissett, 1985; Snyder et al., 1977), and (iv) less competent individuals are less able to take advantage of feedback via social comparison[4].

There is, however, an alternative explanation, particularly in the case of government officials: government officials may not recognize their own shortcomings and try to improve their knowledge or skills because, at least in some cases, they are not incentivized to do so. We should first note that government's meta-ignorance can be costly for society at large. Take, for example, the case of 2008 Global Financial Crisis that originated in United States. Failure of regulators to realize how little they know about then-unknown adverse impacts of unchecked financial innovations made them put too much faith in market's self-regulation, leading to near-total collapse of global financial market. This is despite that the source of crisis, elevated stress on U.S. subprime mortgage market, were not actually large enough to cause crisis in such a global scale by itself.

2.2 Why Remaining Blissfully Ignorant of One's Ignorance Can Be Rational

If both the cost of being ignorant of one's ignorance and the benefit of continuous learning are significant, as mentioned above, and if private incentives of individual government officials are perfectly aligned with social interest, government officials may become incentivized to remain vigilant of the "unknown-unknown" by accumulating more knowledge. Accumulating more knowledge allows individual officials to provide more social benefits, which also benefits him/her. Therefore, in the ideal world, officials who serves the public will wisely try to pursue public interests by being aware of their own shortcomings and thus intrinsically incentivized to learn more.

However, as in many other cases involving public sector, there is principal-agent issue in the problem of government officials' meta-ignorance; while mistake is costly for taxpayer, private cost for government officials are negligible. Likewise, officials do not receive pecuniary compensation for good services, which can only be made possible by accumulating sufficient knowledge. This paper makes the case that under certain condition, incentive for government officials to be aware of the extent and limit of their knowledge may become negligible and makes them less likely to learn more, which in turn makes them unable to better recognize the extent of deficit of their knowledge (i.e. ignorant of one's ignorance can be individually rational).

This paper incorporates insights from contract theory by utilizing basic framework à la Laffont and Martimort (2002) and Delfgaauw & Dur (2008) to explain why government officials with low level of prior knowledge tends to overestimate their knowledgeability and thus learn less, while those with more prior knowledge tends to learn more

Let us assume that government, which for simplicity is assumed to be representation of the people at large[5], acts

[3] False consensus effect in this context refers to perception among people who are knowledgeable and/or top performer to falsely assume that since they are performing so well (perhaps effortlessly), they assume that their peers should also perform as well as them. This issue is also discussed in Kruger & Dunning (1999).

[4] In the words of Kruger & Dunning (1999), less competent individuals are less able to spot competence when they see it compared to their more competent counterparts, so that watching behavior do not make their estimate of their own ability is incorrect.

[5] We also simplify the model by assuming that government office's incentive is aligned with community at large, i.e. what government office wants is what the community want.

as principal and individual officials as its agents. Government provides benefit for society to the value of $S(\sum_{t=1}^{T}(1+\delta)^{t-T}q_{i,t})$, which can be simplified to $S(\cdot)_{i,t}$. Stripped to its essence, the work of government is to mobilize officials to use their knowledge and expertise to provide services to general public, so that social value of government services $S(\cdot)_(i,t)$ depends on officials' accumulated knowledge $(\sum_{t=1}^{T}(1+\delta)^{t-T}q_{i,t})$. Social benefit is assumed to follow:

$$S'(\cdot)_{i,t} > 0 \tag{2.1a}$$

$$S''(\cdot)_{i,t} < 0 \tag{2.1b}$$

$$S(0) = 0 \tag{2.1c}$$

Accumulation of knowledge by government officials affect quality of public services $S(\cdot)_{i,t}$ through (i) reduction of expected social cost from the possibility of avoidable, costly mistakes and (ii) improved quality of government services delivery. To illustrate the two, let us use the example of public planning within the context of waste management. Knowledge about increased risks of disease outbreak and flooding from allowing households to dump their waste into the river allows officials to take preventive measures, such as by routinely collecting household wastes and promoting privately-operated recycling center to reduce the cost. Additionally, preventive better waste management system reduces environmental nuisances, such as terrible smell and unsightly floating trashes, and improve public satisfaction from better water quality.

Individual knowledge of a given subject is assumed to decay over time, a concept reflected by the term δ. Knowledge decay happens naturally through memory loss of any given knowledge over extended period of time. Ability to recollect information from texts that were read 10 years ago, for example, is lower than from texts read just yesterday. Information that was acquired from distant past is also less useful than, which further contributes to effective decay of knowledge over time. The term δ is close to the concept of depreciation, but instead of depreciation of physical capital stock, δ represents depreciation of stock of knowledge; this phenomenon is also more popularly known as half-life of knowledge.

For individuals, the process of learning itself carries two implications. First, accumulation of knowledge allows individual to know better the extent of total available knowledge to learn. Someone who only took Economics 101 may readily think that everything in the introductory textbook is all there is to learn about economics, whereas someone who have obtained Ph.D. in Economics will be acutely aware of the breadth of the field of economics to the point where he or she may conclude that they can only know in depth about a tiny subset of economics. We can represent one's awareness about the breadth and depth of a given field of knowledge with $Q_{i,t}$, where:

$$dQ_{i,t}/d\left(\sum_{t=1}^{T}(1+\delta)^{t-T}q_{i,t}\right) > 0 \tag{2.2a}$$

$$d^2Q_{i,t}/d\left(\sum_{t=1}^{T}(1+\delta)^{t-T}q_{i,t}\right)^2 \geq 0 \tag{2.2b}$$

The consequence of (2.2a) and (2.2b) is that theoretically, individuals will realize that it is impossible to run out of new knowledge to learn, even (or, perhaps, particularly) when someone has become an expert in a given field.

Second, the outcome of learning process generates utility for individual officials, thus making accumulation of knowledge. Individual officials, being a part of society, stands to benefit from the social benefit generated by their own work, $S(\cdot)_{i,t}$. The utility derived from their own work is represented by $\gamma_{i,t}S(\cdot)_{i,t}$, where:

$$\gamma_i = f\left(\sum_{t=1}^{T-1}(1+\delta)^{t-T}q_{i,t}\right) \in [0,1] \tag{2.3a}$$

$$d\gamma_i/d\left(\sum_{t=1}^{T-1}(1+\delta)^{t-T}q_{i,}\right) > 0 \tag{2.3b}$$

$$d^2\gamma_i/d\left(\sum_{t=1}^{T-1}(1+\delta)^{t-T}q_{i,t}\right)^2 < 0 \tag{2.3c}$$

Intuitively, γ_i indicates that for a given level of $S(\cdot)_{i,t}$, the intrinsic utility (i.e. how much they appreciate the social benefit that comes from their work) increases if their existing knowledge increases. The concept of $\gamma_{i,t}$ can be demonstrated using an example about waste management. If our hypothetical official is not aware of the risks associated with dumping household waste into the river (no prior knowledge on the subject of waste treatment, $\sum_{t=1}^{T-1}(1+\delta)^{t-T}q_{i,t} = 0$), s/he will put little to no value ($\gamma = 0$) on learning waste management best practice. S/he will also have no problem with continuous dumping of household waste into the river or letting the community operate open landfill in densely-populated area, thus risking disease outbreak, flooding and reduction in overall environmental quality.

The process of learning is, however, privately costly. Individual officials paying for both fixed[6] (F) and variable cost (θ) of knowledge acquisition. For simplification, this paper sets the fixed cost to zero. Individual officials also face different variable costs of acquiring knowledge; highly productive officials (denoted by superscript H) can learn easily and face lower cost of learning than less productive officials (denoted by superscript L), thus implying $\theta_H < \theta_L$. This implies that the cost structure for individual officials equals to:

$$C(q,\theta_H) = \theta_H q_{H,t} \tag{2.4a}$$

$$C(q,\theta_L) = \theta_L q_{L,t} \tag{2.4b}$$

[6]This fixed cost is also known as status quo utility level, i.e. reservation utility level in the form of outside opportunity that can be obtained outside this transaction.

Officials type i derive benefit from their respective wage (w) and aforementioned intrinsic utility from learning, which is represented by the following function:

$$U_{i,t} = w_{i,t} + \gamma_i S(\cdot)_{i,t} - \theta_i q_{i,t} \tag{2.5}$$

Before entering a contract, officials (agents) are assumed to have ex-ante knowledge of their own type. The government (principal) also have knowledge about general distribution of each type of worker, but cannot assess the type of each individual worker. The problem faced by the principal is, therefore, to maximize the following program, subject to participation and incentive constraints of both highly productive and less productive officials:

$$max(p\Big(S(\cdot)_{H,t} - w_H\Big) + (1-p)\Big(S(\cdot)_{H,t} - w_L\Big)) \tag{2.6}$$

In the first-best world, government as principal can offer contract that yields zero utility for each and every agent ($w_{i,t} = \theta_i q_{i,t} - \gamma_{i,t} S(\cdot)_{i,t}$) in order to achieve its own best possible utility. The complication to this first-best scenario arises from inability of government as principal to correctly assess the type of each individual worker and incentive constraint of highly productive worker. As highly productive officials have lower learning cost than less productive officials, they can gain positive utility by mimicking less productive worker; by definition, if government can offer less productive workers a contract of $w_L - \theta_L q_{L,t} + \gamma_L S(\cdot)_{L,t} = 0$, highly productive worker can take that contract and have utility of $(\theta_L - \theta_H) q_{L,t} + (\gamma_{H,t} - \gamma_{L,t}) S(\cdot)_{H,t} > 0$. Thus, if the government as principal wants highly productive officials to elicit the private information regarding their true type, the contract offered to them have to satisfy the incentive constraint as outlined in (2.7), which can be re-written as (2.8) or (2.9)

$$w_{H,t} + \gamma_{H,t} S(\cdot)_{H,t} - \theta_H q_{H,t} \geq w_{L,t} + \gamma_{H,t} S(\cdot)_{L,t} - \theta_H q_{L,t} \tag{2.7}$$

$$U_{H,t} \geq U_{L,t} + (\theta_L - \theta_H) q_{L,t} + (\gamma_{H,t} - \gamma_{L,t}) S(\cdot)_{L,t} \tag{2.8a}$$

$$U_{H,t} \geq U_L + \Delta\theta q_{L,t} + \Delta\gamma_t S(\cdot)_{L,t} \tag{2.8b}$$

$$w_{H,t} - w_{L,t} \geq \theta_H (q_{H,t} - q_{L,t}) - \gamma_{H,t} (S(\cdot)_{H,t} - S(\cdot)_{L,t}) \tag{2.8c}$$

$$\Delta w_t \geq \theta_H \Delta q_t - \gamma_{H,t} \Delta S(\cdot)_t \tag{2.8d}$$

The reverse, however, can be safely ignored. As the cost of carrying out the contract for less productive officials are higher than for highly productive officials, the contract offered to highly productive official is unappealing to less productive officials. This condition has two implications. First, if a menu of contracts can satisfy less productive officials' participation constraint ($U_{L,t} \geq 0$), that menu will always strictly satisfy highly productive agent's participation constraint. The constraints for principal's optimization can therefore be simplified to less productive officials' participation constraint and highly productive officials' incentive constraint (Laffont & Martimort, 2002:Chapter 2).

Rewriting highly productive officials' incentive constraint (2.9b) and less productive officials' participation constraint into principal's problem in (2.6) allows us to have outputs as the problem's only choice variable:

$$\begin{aligned} max(&p\Big(S(\cdot)_{H,t} - \theta_H q_{H,t} + \gamma_{H,t} S(\cdot)_{H,t}\Big) \\ &+ (1-p)\Big(S(\cdot)_{L,t} - \theta_L q_{L,t} + \gamma_{L,t} S(\cdot)_{L,t}\Big) \\ &- p\Big(\Delta\theta q_{L,t} + \Delta\gamma_t S(\cdot)_{L,t}\Big)) \end{aligned} \tag{2.9}$$

Maximization of principal's utility with respect to amount of knowledge learned by officials at period T ($q_{i,T}$) yields:

$$S'(\cdot)_{H,T} = \frac{1}{1+\gamma_H} \theta_H < \theta_H \tag{2.10}$$

$$S'(\cdot)_{L,T} = \frac{(1-p)\theta_L + p\Delta\theta}{(1-p)(1+\gamma_{L,T}) - p\Delta\gamma_T} > \theta_L \tag{2.11}$$

(2.10) and (2.11) shows that the second-best, incentive-feasible menu of contracts entails upward distortion of the amount of knowledge acquired at period T by highly productive officials and downward distortion of the amount of knowledge acquired at period T by less productive workers when compared to first-best output, consistent with Delfgaauw & Dur (2008). At this point, let us introduce a special case where less productive workers also have zero prior knowledge, so that $\gamma_{L,T} = 0$. This brings (2.11) to:

$$S'(\cdot)_{L,T} = \frac{(1-p)\theta_L + p\Delta\theta}{1 - p(1+\gamma_{H,T})} \tag{2.12}$$

(2.12) represents further downward distortion in output that happens if less productive officials, due to lack of knowledge, do not appreciate the social benefit created by accumulating more knowledge and do not derive any intrinsic utility from their own work. Using another example, officials that have little to no knowledge about the link between human activities climate change may not see the intrinsic benefit of learning more on subjects like climate change mitigation or adaptation. Less productive officials will therefore only commit to learn the bare minimum level required by their respective job description (only meeting the participation constraint). By comparing (2.12) with (2.11), we can infer that the practice of routinely rotating individual officials from one function to another, which does not allow them to specialize in a specific function, can be counterproductive and socially undesirable.

The more sanguine way to look at (2.11) and (2.12) is that as less productive officials start to accumulate more knowledge, they are also more appreciative of the benefit of acquiring more knowledge, particularly the limitation of their own knowledge, thus increasing the amount of knowledge they acquire with same level of wage. This means that incentivizing officials to learn may, in the long run, increase their accumulated level of knowledge in that subject and

allows them to accurately assess their level of knowledge (i.e. by realizing that there is always more to learn), despite their initial objections to do so.

By looking at the incentive-compatible level of q_H^* and q_L^*, we can inspect how each individual assess their level of knowledge ($\sum_{t=1}^{T}(1+\delta)^{t-T}q_{i,t}$) when compared to their perceived breadth of knowledge available to learn in a given field ($Q_{i,t}$). Since $q_H^* > q_L^*$ at the incentive-compatible optimum level, we can infer from (2.2b) and (2.3b) that:

$$\frac{\sum_{t=1}^{T}(1+\delta)^{t-T}q_H^*}{Q_H^*} < \frac{\sum_{t=1}^{T}(1+\delta)^{t-T}q}{Q_L^*} \tag{2.13}$$

(2.13) highlights how, at their respective optima, the knowledge accumulated by both highly productive officials and less productive officials is such that highly productive (and thus knowledgeable) officials think that they only know relatively little compared to the vastness of knowledge that is available to learn, while less productive (and thus less knowledgeable) officials are more confident with their arguably less knowledge. Therefore, (2.13) exhibits the phenomenon of Dunning-Kruger effect. It should be stressed that for both types of individuals, the level of accumulated knowledge satisfy their individual optimization, thus making their ignorance a rational one.

The second-best, incentive-compatible contracts as shown by (2.10) and (2.11), also highlight the problem with existing remuneration scheme offered by government to its officials. Ideally, highly productive officials should be demanded to focus their energy on gaining more knowledge and learning new, more innovative approaches to government's day-to-day problems and thus rewarded accordingly. However, by offering identical wage for both types of officials, as is the case today, highly productive workers will only learn more to the extent of their perceived utility from learning more, which may or may not amount to much. Under the current condition, highly productive officials are willing to accept identical wages while learning more than less productive officials by

$$\Delta q = \frac{\gamma_{H,T}(S(\cdot)_{H,T} - S(\cdot)_{L,T}^*)}{\theta_H} \tag{2.14}$$

Viewed from another perspective, offering contracts with similar wage as in (2.14) is akin to purely appealing to altruistic motive when it comes to pushing highly productive officials to learn more. Compare (2.14) with difference in quantity of learning between highly productive and less productive workers under second-best, incentive-compatible contracts (i.e., when paid sufficiently more):

$$\Delta q^* == \frac{(w_{H,T}^* - w_{L,T}^*) + \gamma_{H,T}(S(\cdot)_{H,T}^* - S(\cdot)_{L,T}^*)}{\theta_H} \tag{2.15}$$

3. METHODOLOGY

3.1 Data

Data for this research is sourced from LPEM FEB UI's Survey on Readiness of Local Government for Implementation of Green Budgeting. The survey was aimed at senior civil servants in local government offices (*Satuan Kerja Perangkat Daerah/SKPD*) who are in charge of climate-change related issues and/or formulation of contribution of local office in Regional Action Plan for Greenhouse Gas Emission Reduction (RAD-GRK). Respondents should be either Heads of Local Office (*Kepala SKPD*), Secretaries of Local Office (*Sekretaris SKPD*, equivalent to Secretary-General in title), Division Heads (*Kepala Bidang*), or Subdivision Heads (*Kepala Sub-Bidang*), mentioned according to level of seniority. Respondents are composed of 138 local offices in 4 provinces and 13 regencies in Indonesia

3.2 Method

The survey is designed to reflect self-assessments on various aspects of readiness to implement green budgeting. Several questions are designed as pairs of "reference question" and three "test questions". Reference question is self-assessment on their ability to conduct certain activities, with responses rated on an ordinal scale of 1–4. Test questions are binary response (yes/no) questions, the responses of which reflect subset of knowledge that are inquired in Reference Question. The aim of this design is to test the consistency of self-assessment that is reported in reference questions and their response, i.e. whether the claim to knowledge that is reported in reference question is warranted.

There are four pairs of Reference Question and their then compared to aggregate of its three respective questions. To do so, the binary responses to Test Questions will be transformed into 1-4 scale that is used for Reference Questions. We start with score of 1, and add score of 1 for each "yes" answer to binary response question. With 3 Test Questions to be compared with each Reference Question, the minimum score for the aggregated Test Questions is 1 (i.e. when the response to every test question is "no" and maximum score is 4 (i.e. when the response to every test question is "yes"), identical with the ordinal scale employed by Reference Question. The list of Reference Questions and Test Questions is provided in the appendix.

Even as the question purportedly try to assess the competence of the whole local office not the competence of the individual respondent in conducting certain activities, the construction of the questions is such that it will reveal the extent of knowledge about concepts that were subject of inquiry in reference questions. For example, in order to accurately assess the ability of local office to prepare RAD-GRK document in Reference Question 1, respondent should know the processes of drafting and preparing RAD-GRK, which includes but is not limited to coordination meeting on RAD-GRK with other local offices (Test Question 1), providing necessary data (Test Question 2), and actually preparing RAD-GRK document (Test Question 3). Accurate metacognitive knowledge should result in consistent result between Reference Questions and their respective Test Questions (i.e. difference between the score of Reference Question and aggregated Test Questions is zero or close to zero).

To detect whether Dunning-Kruger effect is present within government circle, we then test the hypothesis of whether the average of self-assessed responses to Reference Questions is equal to aggregate of response to their respective Test Questions (using simple t-test), and plot the results into scatterplot. Ordered probit regression is then performed

Figure 1. Response to Statement "This Local Office is Able to Prepare the Regional Action Plan for Greenhouse Gas Emission Reduction (RAD-GRK) document"

to detect the source of overestimation in self-assessment on selected aspects of green budgeting readiness asked on Reference Questions.

4. RESULTS AND DISCUSSIONS

We compare the response for reference questions, which serve as respondent's self-assessment, and their respective questions, which serve as respondent's actual knowledge. We illustrate this by plotting respondent's actual knowledge against respondent's self-assessment in one scatterplot. The result, presented below, suggests that government officials who do not have much knowledge about aspects of green budgeting tend to overestimate their abilities, while those who perform highly in actual measurement tend to underestimate their abilities. This result is in line with findings of Kruger & Dunning (1999) and Ehrlinger et al. (2008).

We first analyze the comparison between self-assessment and actual result (derived from its three constituent Test Questions) for Reference Question 1. Overall, self-assessment for Reference Question 1 ($\bar{x} = 2.6769$) is higher than the actual result ($\bar{x} = 2.3615$), $t(129) = 3.0436$, p<0.01. The relationship between self-assessment and actual result is also only weakly positive and statistically insignificant, with $r_s = 0.1210$, $p = 0.1704$.

Next, we analyze the comparison between self-assessment and actual result (derived from its three constituent Test Questions) for Reference Question 2. Overall, self-assessment for Reference Question 2 ($\bar{x} = 2.4253$) too is higher than the actual result ($\bar{x} = 1.7537$), $t(133) = 7.7587$, p<0.01. The relationship between self-assessment and actual result is moderately positive and statistically significant, with $r_s = 0.4241$, $p = 0.0000$.

It should be noted that difference between stated ability to calculate level of emissions and actual ability (proven when asked about the ability to calculate emissions related to their activities and whether they have done so in the past) are more pronounced than in other Reference Questions. This can be explained by the fact that most respondents have little awareness that activities related to their field can have contribution and thus are never aware of how estimation of greenhouse gas emission actually works. More respondents thus tend not to realize what process of estimation of greenhouse gas emission entails.

We then analyze the comparison between self-assessment and actual result (derived from its three constituent Test Questions) for Reference Question 3. Overall, self-assessment for Reference Question 3 ($\bar{x} = 2.5426$) is higher than the actual result ($\bar{x} = 2.1550$), $t(128) = 4.2611$, p<0.01. The relationship between self-assessment and actual result is moderately positive and statistically significant, with $r_s = 0.4523$, $p = 0.0000$.

We finally analyze the comparison between self-assessment and actual result (derived from its three constituent Test Questions) for Reference Question 4. Overall, self-assessment for Reference Question 4 ($\bar{x} = 2.7272$) is higher than the actual result ($\bar{x} = 2.3561$), $t(131) = 3.5907$, p<0.01. The relationship between self-assessment and actual result is also weakly positive and statistically insignificant, with $r_s = 0.1229$, $p = 0.1605$.

The range of averages of the self-assessment for Reference Questions is 2.43–2.73, which is higher and has less range than the average results of actual assessment (Range 1.75–2.36). While average responses to Reference Question 1, 2, 3, or 4 are not statistically the same, with $F(539) = 6.12$, $p < 0.01$, breakdown of descriptive statistics shows that responses to Reference Questions are centered around the score of 2 and 3, as seen at the table below. This also suggests that government officials also tend to rate themselves as "above average", in line with results of previous studies.

We then try to assess what factors are likely to affect officials to be self-confident by running. As descriptive statistics show that most of the self-assessment scores is 3 on the scale of 1–4, we will regress both responses to Reference Questions and Test Question with ordered probit model, focusing the result on probability of someone to score 3 on the scale of 1–4.

Our findings show that the overall pattern tend to be very weak (models' R2 below 0.10) yet statistically significant. The clearest tend that emerge is that respondents from Office of Environmental Affairs (*Badan Lingkungan Hidup Daerah*), on average, is more likely to underestimate their knowledge compared. This trend emerges when we compare marginal effects of Reference Question 1 (self-

Figure 2. Response to Statement "This Local Office is Able to Calculate Level of Emissions and Emission Reduction from Mitigation Activities in Every Sectors that are Under the Purview of Local Office"

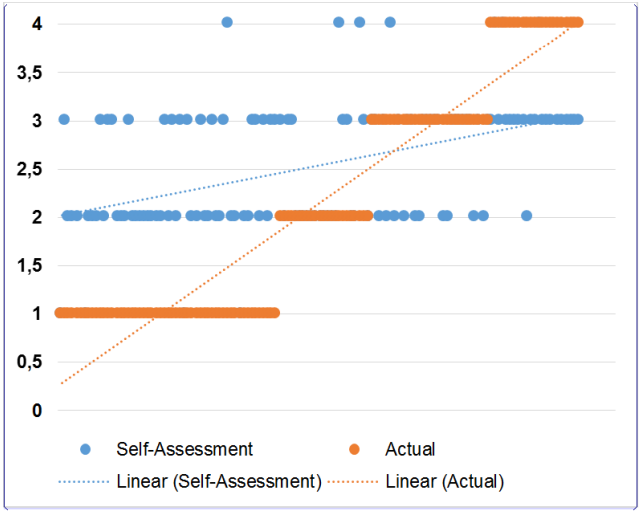

Figure 3. Response to Statement "This Local Office is Able to Prepare the Monitoring, Evaluation, and Reporting of Regional Action Plan for Greenhouse Gas Emission Reduction (RAD-GRK) document"

assessment) to Test Questions 1 (actual results) and Reference Question 2 to Test Questions 2. In Question 1, offices of environmental affairs are 1.97% (n.s.) more likely to score 3 on self-assessment, but 7.23% ($p<0.10$) more likely to score 3 on actual results. In Question 2, offices of environmental affairs are 6.36% (n.s.) more likely to score 3 on self-assessment, but 9.14% ($p<0.05$) more likely to score 3 on actual results. The pattern in 2 other questions are not clear and not statistically significant.

This trend may be attributed to the fact that offices of environmental affairs are most likely to have issues of climate change and green budgeting under their purview, thus making officials there to be more aware of steps of climate change mitigation and of green budgeting, and more compliant with RAD-GRK. This awareness of issues related to green budgeting and climate change mitigation makes officials in office of environmental affairs more likely to accurately assess their knowledge of issues in question.

Another relatively clear trend is consistently significant overestimation among government officials in Nusa Tenggara Timur, particularly when compared to other provinces. Compared to baseline province (Sulawesi Barat) in Question 1, local officials in Nusa Tenggara Timur, both province-level and regency level, are 16.43% ($p<0.05$) more likely to score 3 on self-assessment, but 1.16% (n.s.) less likely to score 3 on actual results. In Question 2, local officials in Nusa Tenggara Timur are 26.05% ($p<0.01$) more likely to score 3 on self-assessment, but only 1.10% (n.s.) more likely to score 3 on actual results. In Question 1, local officials in Nusa Tenggara Timur are 8.37% (n.s.) more likely to score 3 on self-assessment, but 2.25% (n.s.) less likely to score 3 on actual results. The same trend plays out in Question 4; local officials in Nusa Tenggara Timur are 21.96% ($p<0.01$) more likely to score 3 on self-assessment, but only 9.28% ($p<0.05$) more likely to score 3 on actual results.

There is no conclusive explanation as to why civil servants from local offices in Nusa Tenggara Timur tend to overestimate their ability by a wide margin, particularly compared to other provinces. However, there are two phenomena that may be related to this overestimation. First, it is possible that Insularity of civil servants and socioeconomic condition may be the culprit, as Nusa Tenggara Timur is the poorest province in Indonesia and civil servants function as local elites. This may give heightened illusion of superiority

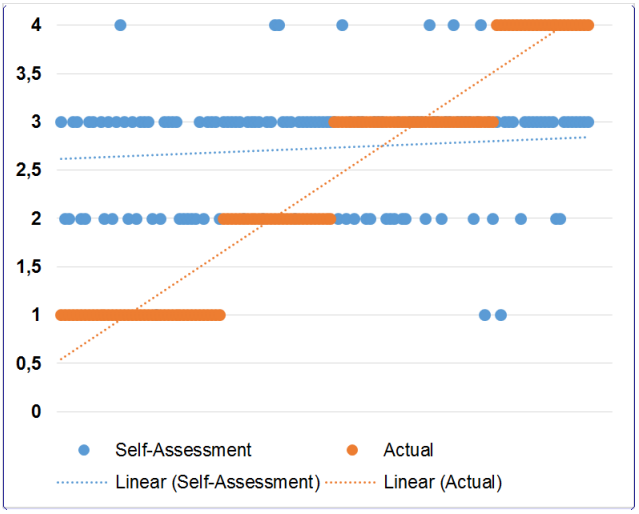

Figure 4. Response to Statement "This Local Office is Able to Incorporate Issues Related to Climate Change within the Planning and Budgeting Cycle"

Table 1. Summary of Responses to Reference Questions (Self-Assessment)

Score	Reference Question 1		Reference Question 2		Reference Question 3		Reference Question 4	
	Freq	Percent	Freq	Percent	Freq	Percent	Freq	Percent
1	4	2.96	4	2.94	4	2.99	3	2.22
2	46	34.07	75	55.15	62	46.27	38	28.15
3	75	55.56	52	38.24	61	45.52	87	64.44
4	10	7.41	5	3.68	7	5.22	7	5.19
Total	135	100.00	136	100.00	134	100.00	135	100.00

among civil servants in NTT relative to other provinces. However, this explanation is not sufficient and can be easily refuted, given that Nusa Tenggara Barat and Sulawesi Barat are also among poorer provinces in Indonesia and local civil servants also serve as local elites, much like their counterparts in NTT.

The second, more likely to explain the trend of overestimation of knowledge among NTT officials, is the existing environmental condition of Nusa Tenggara Timur. Different with most other provinces in Indonesia, major islands in NTT, such as Sumba and Timor, have savannah climate. This naturally drier condition may induce officials and population in general not to attribute adverse climate events, such as drought or flash flood, to human activities, but more as the way nature works. This may also explain why officials in NTT are less inclined to care or learn about human negative impact on environment because they have lived with less friendly environmental conditions for a very long time.

Less apparent but equally interesting regional differences in score can be seen in relatively higher mean test score (actual ability) in Jambi and propensity of respondents in Jambi to be more accurate in measuring their readiness in aspects of green budgeting, compared to other provinces (q.v. Appendix Table A1). While respondents in Jambi still tend to overestimate their readiness in aspects of green budgeting, the gap between self-assessment and actual assessment tend to be much lower than in other provinces. Indeed, regression results suggest that respondents in Jambi might actually tend to slightly underestimate their ability; in Question 1, Jambi officials are 11.53% ($p < 0.10$) more likely to score 3 on self-assessment, but 20.29% (n.s.) more likely to score 3 on actual results. Likewise, in Question 4, Jambi officials are 6.01% (n.s.) more likely to score 3 on self-assessment, but 6.31% ($p < 0.10$) more likely to score 3 on actual results.

The most plausible reason behind why Jambi officials tend to outperform three other provinces in terms of accurately measuring their own readiness on aspects of green budgeting is because Jambi experienced large-scale anthropogenic environmental disaster. In 2015, Jambi experienced total shutdown as forest fires, which were used to clear lands for new palm oil plantations, became uncontrollable. This months-long continuous forest fires, which prevented people in Jambi to do any activities outside homes and claim many casualties, caused deep trauma among people in Jambi, as also shown in interview with both officials and locals. This shock prompted Jambi provincial government and regency/municipal governments to dedicate significantly more attention to environment-related issues, which then explains higher awareness and knowledge of environmental issues across-the-board in Jambi.

A surprising result come from positive statistical relationship between proportions of undergraduate level (S1) employee to total employee (proxy of how well-educated civil servants in each local office are) and overestimation of level of knowledge among respondents. In Question 1, increase of 1 percentage point in proportion of undergraduate-level employee makes respondent 0.20% (n.s.) more likely to score 3 on self-assessment, but only 0.17% ($p < 0.10$) more likely to score 3 on actual results. In Question 2, increase of 1 p.p. in proportion of undergraduate-level employee makes respondent 0.40% ($p < 0.10$) more likely to score 3 on self-assessment, but only 0.21% ($p < 0.05$) more

likely to score 3 on actual results. In Question 3, increase of 1 p.p. in proportion of undergraduate-level employee makes respondent 0.50% ($p < 0.01$) more likely to score 3 on self-assessment, but only 0.14% ($p < 0.05$) more likely to score 3 on actual results. Question 4 yields no statistically significant measure.

The effect is more acute when we measure the portion of postgraduate (S2 and S3) level employee and how it makes respondent more likely to overestimate their knowledge and ability. In Question 3, increase of 1 p.p. in proportion of undergraduate-level employee makes respondent 1.04% ($p < 0.05$) more likely to score 3 on self-assessment, but 0.12% (n.s.) less likely to score 3 on actual results. In Question 4, increase of 1 p.p. in proportion of undergraduate-level employee makes respondent 0.50% ($p < 0.01$) more likely to score 3 on self-assessment, but only 0.14% ($p < 0.05$) more likely to score 3 on actual results. Regression with respect to self-assessment responses and actual results for Question 1 and Question 2 do not yield statistically significant results.

Illusion of superiority may again account for tendency for respondent to overestimate when they know their employees are better-educated. It may very well be the case that when senior civil servants know that their peers and subordinates are well-educated, they become more likely to assume that at least somebody in their offices know about concepts related to green budgeting, even if they do not personally know about the concepts in question. However, more researches are needed to better account for this phenomenon.

4.1 Limitation of This Study

We should note that the survey used by this study relies almost completely on self-assessment; even the Test Questions, which are supposed to measure actual skills, still rely to some degree on self-assessment, as presenting specific question in binary response format may reduce cognitive load needed to answer the questions and are not less accurate than ordinal response format (Dolnicar et al., 2011). The binary response format used in this survey is also designed to reveal skills even when relying on self-assessment, given that most of the questions focus more on whether the respondents (who represent their respective local offices) *have done something* and less on whether the respondents *can do something*. Binary response format is also less prone to above-average effect, given that they do not leave room for people to rate themselves as "not the best but above-average". Furthermore, steps have also been taken to ensure that the question is not asked by surveyor in biased manner.

Nevertheless, without actual testing, we cannot ascertain the degree of accuracy of Test Questions. Responses to Test Questions in themselves may also be biased by Dunning-Kruger effect, particularly for Test Questions that ask respondents about whether they *can do something* (it is almost impossible to present inaccurate responses to questions of whether they *have done something*, unless the respondents are being untruthful). Further researches in this direction should devise quantitative, non-self-assessed survey to accurately assess government performance in any subject.

On the other hand, devising proper examination to assess senior local civil servants' actual performance on green budgeting, or any other government activities, is also impractical, given the long duration it takes to assess competence on certain government activities and time constraint of senior civil servants. For example, measuring ability of government officials to conduct social cost-benefit analysis, even when key data and assumptions are provided, may take hours, given the complexity of calculating social cost and social benefit. With time constraints faced by civil servants, it may be very difficult to find enough willing civil servants to have any statistically meaningful survey.

The very problem of higher user fatigue for ordinal responses format (Dolnicar et al., 2011) may also affect the quality of responses in this survey. Simple, one-stage ordinal survey that is employed in LPEM's survey may also be prone to central tendency error and error in direction sign, which also impact the ability of respondent to answer our reference questions accurately in the first place. That the survey takes each respondent around 30 minutes on average can also affect the accuracy of the responses, as mental fatigue may induce respondents not to pay full attention to questions that are being asked.

4.2 Policy Implication

Should further studies confirm that government officials are themselves prone to Dunning-Kruger effect, even when they are not supposed to, the implication for public governance may be enormous. This highlights the critical need for government officials as individuals and collectively as governing bodies to improve awareness on the "known-unknown" and "unknown-unknown" in their respective fields, particularly as issues that have to be dealt by the government become increasingly complex. We highlight several insights that arise from realization that less knowledgeable government officials are no better in avoiding Dunning-Kruger effect than general population:

1. Whenever government, particularly central government, want to introduce new system, which is exemplified by green budgeting and RAD-GRK in this study, **self-assessment and self-monitoring are grossly inadequate**. Taking the RAD-GRK and green budgeting implementation for local governments as example, **measurement of ability**, performance, and readiness to implement new system **should be conducted by external parties**.
2. **Good performance measurement design need to be easily quantifiable** (so that measurements are comparable) **and be as objective as possible**. Easily quantifiable and transparent measurement of officials' individual performance and collective departmental performance helps to incentivize officials to formulate clear and workable plan to achieve the pre-set target.
3. **Large scale adverse events, may tend to promote learning and more accurate self-assessment**. Adverse events, such as large-scale anthropogenic forest fire experienced by Jambi and other provinces in Sumatera, tend to promote awareness of that topics for those who are unaware and is one of the most effective way for government and individuals alike to learn and better understand the limit of one's ignorance. Similarly, it is only after large scale economic

crisis in 1997–98 that Bank Indonesia fully recognize the risks and true costs of adopting crawling peg and went on to float Rupiah ever since. Negative shocks, either through ignorance, mistakes, or bad luck, are the best teachers, particularly when such events are privately costly for decision makers.

Good incentives are critical in nudging government officials to improve their metacognitive skills. Good designs should incentivize government officials as individuals and as a group to become fully aware of the limit of their knowledge and to improve their knowledge so they can perform their jobs better. Taking the first and second insights, first step towards better governance is to set realistic, quantifiable performance measurement and target for local government. As a practical suggestion with respect to green budgeting, Ministry of Interior may want to consider implementing clear target and timeline for local government offices to implement green budgeting. This can be done with simple target, such as 'requiring 20% reduction in total carbon emissions from government activities by 2020'.

Using the third insight, good incentives should give signal to government officials that ignorance, and mistakes resulting from it, are privately costly and that good performance will be well-rewarded. The key takeaway from our theoretical model regarding wage differential, be it in the form of expected future promotion (present value of expected incremental future income) or explicitly paying high performer with better salary, is that it should be tied to learning-related performance ($\theta_H(q_{H,t} - q_{L,t})$).

Additionally, as previously discussed, given that less productive workers do accumulate their knowledge over time, it might not be in the best interest of government to frequently rotate officials to different functions. Whenever local officials are transferred to new office following change in leadership, officials will start from scratch with respect to knowledge in their new function, thus setting the γ_L to zero again. This will reduce overall efficiency of the government, as they can only expect officials to learn less for the same wage whenever they are moved to different functions.

One of the more practical suggestion on this with respect to green budgeting is for local offices to name officials in charge of emission reduction target, with clear risk and reward, both for ability to accurately measure CO_2 reduction (requiring knowledge of GHG emission calculation) and for ability to meet the target in time (requiring knowledge of cost-benefit analysis). Before emission reduction program is implemented, independently-conducted pre-test, training, and post-test to measure initial readiness of local offices (particularly for knowledge of GHG emission calculation and cost-benefit analysis) should be held. There should also be routine target monitoring to measure accuracy of GHG emission measurement by local offices and compare actual reduction with target reduction. Ultimately, the government should remove

Aligning private incentives of government officials and public interest is not only useful in the context of green budgeting, but can also be implemented in other areas of public governance. We should however always be aware that current government structure, with ineffective legislative bodies and government departments that tend to self-regulate, may present difficulties for realignment of civil servants' incentives with social interest. To reiterate the point of Regehr & Eva (2006), people who are aware of their shortcomings people may still avoid learning when such learning takes more energy and commitment than they are willing to expend.

5. CONCLUDING REMARKS

Results of this study on green budgeting readiness give strong indication that government officials are not immune to Dunning-Kruger effect. Further analysis reveals that certain government function, such as respondents from office of environmental affairs, have better metacognitive skills when it comes to concepts of green budgeting, given that they deal with climate change mitigation-related issues on daily basis. Regions that are more. Presence of more educated workforce, paradoxically, give illusion of superiority amongst respondents and cause respondent to overestimate their own office's ability.

This study yields three key insights from the presence of Dunning-Kruger effects in government officials circle, particularly among the less-knowledgeable. First, self-assessment and self-monitoring are grossly inadequate, which suggest that if central government wants local governments to implement green budgeting, they should create mechanism to independently assess the readiness and performance of local government. Second, good performance measurement design should be easily quantifiable and objective; targets for green budgeting for local governments cannot be vague or unmeasurable. Third, large-scale adverse events tend to promote learning and more accurate self-assessment, particularly when continued ignorance is costly for decision makers.

This problem of incentives, or lack thereof, may also account for apparent lack of efforts to improve metacognitive ability among civil servants under current setting, as even people who are aware of their shortcomings people may still avoid learning when such learning takes more energy and commitment than they are willing to expend. When ignorance is costly, government officials cannot afford to remain ignorant of their ignorance and shortcomings. The most important part of good incentives for government officials to implement green budgeting (or other policies) is therefore overwhelmingly simple; name official(s) responsible for implementation of green budgeting and ensure that good performances are well-rewarded and bad performances are privately costly.

For full disclosure, errors that may become apparent to future readers are not apparent to author without the benefit of hindsight. Not even the author, or anyone for that matter, is able to be fully free from Dunning-Kruger effect.

References

Brissett, D. (1985). "Reviewed Work: Excuses: Masquerades in Search of Grace' by C. R. Snyder, R. L. Higgins, R. J. Stucky. *Contemporary Sociology*, 14(3), 402–404. Retrieved from http://www.jstor.org/stable/2071382. doi: http://dx.doi.org/10.2307/2071382.

Cross, D. R., & Paris, S. G. (1988). Developmental and instructional analyses of children's metacognition and reading compre-

hension. *Journal of Educational Psychology*, 80(2), 131–142. doi: http://dx.doi.org/10.1037/0022-0663.80.2.131.

Delfgaauw, J. & Dur, R. (2007). Incentives and Workers' Motivation in the Public Sector. *The Economic Journal*, 118(525), pp. 171–191. doi: 10.1111/j.1468-0297.2007.02108.x.

Dolnicar, S., Grün, B., & Leisch, F. (2011). Quick, simple and reliable: Forced binary survey questions. *International Journal of Market Research*, 53(2), 233–254.

Ehrlinger, J., Johnson, K., Banner, M., Dunning, D., & Kruger, J. (2008). Why the unskilled are unaware: Further explorations of (absent) self-insight among the incompetent. *Organ Behav Hum Decis Process*, 105(1), 98–121. doi:10.1016/j.obhdp.2007.05.002.

Eva, K., & Regehr, G. (2005). Self-assessment in the health professions: A Reformulation and research agenda. *Academic Medicine*, 80(Supplement), S46–S54. http://dx.doi.org/10.1097/00001888-200510001-00015.

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. *American Psychologist*, 34(10), 906–911. doi: http://dx.doi.org/10.1037/0003-066X.34.10.906.

Huang, S. (2013). When peers are not peers and don't know it: The Dunning-Kruger effect and self-fulfilling prophecy in peer-review. *Bioessays*, 35(5), 414–416. doi: 10.1002/bies.201200182.

Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: how difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology*, 77(6), 1121–1134. doi: http://dx.doi.org/10.1037/0022-3514.77.6.1121.

Kuhn, D., & Dean, Jr, D. (2004). Metacognition: A bridge between cognitive psychology and educational practice. *Theory Into Practice*, 43(4), 268–273. doi: http://dx.doi.org/10.1207/s15430421tip4304˙4.

Laffont, J. J., & Martimort, D. (2002). *The theory of incentives*. Princeton.

Lai, E. R. (2011). Metacognition: A literature review. *Research report*. Always learning - Pearson. http://images.pearsonassessments.com/images/tmrs/metacognition_literature_review_final.pdf.

Papaleontiou-Louca, E. (2008). M̂etacognition and theory of mind. Newcastle, UK: Cambridge Scholars Publishing. http://www.cambridgescholars.com/download/sample/59586.

Paris, S. G., & Winograd, P. (1990). Promoting Metacognition and Motivation of Exceptional Children. *Remedial And Special Education*, 11(6), 7–15. doi: 10.1177/074193259001100604.

Regehr, G., & Eva, K. (2006). Self-assessment, self-direction, and the self-regulating professional. *Clinical Orthopaedics and Related Research*, 449, 34–38. doi: 10.1097/01.blo.0000224027.85732.b2.

Ross, L., Greene, D., & House, P. (1977). The "false consensus effect": An egocentric bias in social perception and attribution processes. *Journal of Experimental Social Psychology*, 13(3), 279–301. doi: https://doi.org/10.1016/0022-1031(77)90049-X.

Schraw, G., & Moshman, D. (1995). Metacognitive theories. *Educational Psychology Review*, 7(4), 351–371. doi: https://doi.org/10.1007/BF02212307.

Schraw, G., Crippen, K. J., & Hartley, K. (2006). Promoting self-regulation in science education: Metacognition as part of a broader perspective on learning. *Research in Science Education*, 36(1–2), 111–139. doi: https://doi.org/10.1007/s11165-005-3917-8.

Snyder, C. R., Shenkel, R. J., & Lowery, C. R. (1977). Acceptance of personality interpretations: the" Barnum Effect" and beyond. *Journal of Consulting and Clinical Psychology*, 45(1), 104–114. doi: http://dx.doi.org/10.1037//0022-006x.45.1.104.

Sternberg, R. J. (1984). What should intelligence tests test? Implications of a triarchic theory of intelligence for intelligence testing. *Educational Researcher*, 13(1), 5–15. https://doi.org/10.3102/0013189X013001005.

Whitebread, D., Coltman, P., Pasternak, D. P., Sangster, C., Grau, V., Bingham, S., Almeqdad, Q., & Demetriou, D. (2009). The development of two observational tools for assessing metacognition and self-regulated learning in young children. *Metacognition and Learning*, 4(1), 63–85. doi: https://doi.org/10.1007/s11409-008-9033-1.

APPENDIX

Table A1. List of Reference Questions and Test Questions

Reference Question	Test Questions
1. This local office is able to prepare the Regional Action Plan for Greenhouse Gas Emission Reduction (RAD-GRK) document	1. Did this local office, in the last 3 years, attend dissemination seminar(s) and/or coordination meeting(s) on preparation of Regional Action Plan for Greenhouse Gas Emission Reduction (RAD-GRK)? 2. Did this local office, in the last 3 years, provide necessary data for preparation of Regional Action Plan for Greenhouse Gas Emission Reduction (RAD-GRK)? 3. Did this local office, in the last 3 years, prepare Regional Action Plan for Greenhouse Gas Emission Reduction (RAD-GRK)?
2. This local office is able to calculate level of emission reduction from mitigation activities in every sectors that are under the purview of local office	1. Can this local office estimate the emission level from their activities? 2. Can this local office estimate the emission level from activities of the general public? 3. Did this local office, in the last 3 years, calculate level of emissions and changes in level of emissions?
3. This local office is able to prepare the Monitoring, Evaluation, and Reporting of Regional Action Plan for Greenhouse Gas Emission Reduction (RAD-GRK) document	1. Did this local office, in the last 3 years, attend dissemination seminar(s) and/or coordination meeting(s) for preparation of Monitoring, Evaluation, and Reporting of Regional Action Plan for Greenhouse Gas Emission Reduction (RAD-GRK)? 2. Did this local office, in the last 3 years, provide necessary data for preparation of Monitoring, Evaluation, and Reporting of Regional Action Plan for Greenhouse Gas Emission Reduction (RAD-GRK)? 3. Did this local office, in the last 3 years, prepare Regional Action Plan for Greenhouse Gas Emission Reduction (RAD-GRK)?
4. This local office is able to incorporate issues related to climate change within the planning and budgeting cycle	1. Can this local office conduct social cost-benefit analysis for their programs/activities? 2. Can this local office assign monetary value to non-monetary cost/benefit? 3. Can this local office assign monetary value to negative impacts of greenhouse gases emissions?

Table A2. Summary Statistics

Provinces	Question 1		Question 2		Question 3		Question 4	
	Reference	Test	Reference	Test	Reference	Test	Reference	Test
Jambi	2.667***	2.571***	2.524***	2.024***	2.571***	2.357***	2.643***	2.476***
	(0.106)	(0.167)	(0.0978)	(0.185)	(0.114)	(0.170)	(0.0890)	(0.161)
Nusa Tenggara Barat	2.609***	2.087***	2.304***	1.478***	2.435***	1.826***	2.652***	1.870***
	(0.122)	(0.226)	(0.117)	(0.165)	(0.123)	(0.232)	(0.119)	(0.211)
Nusa Tenggara Timur	2.784***	2.216***	2.595***	1.676***	2.595***	2.027***	2.892***	2.757***
	(0.0878)	(0.182)	(0.113)	(0.174)	(0.0985)	(0.196)	(0.101)	(0.171)
Sulawesi Barat	2.524***	2.476***	2.238***	1.714***	2.619***	2.333***	2.571***	2.095***
	(0.164)	(0.235)	(0.118)	(0.250)	(0.146)	(0.279)	(0.148)	(0.248)
Observations	123	123	123	123	123	123	123	123

Standard errors in parentheses *** p<0.01, ** p< 0.05, * p<0.1

Table A3. Ordered Probit Model
Average Marginal Effect, Pr(Ref1 = 3)
Baseline Province: Sulawesi Barat

Independent Variables	Question 1		Question 2		Question 3		Question 4	
	Reference	Test	Reference	Test	Reference	Test	Reference	Test
D˙Pemprov	0.0224	0.5731**	0.1148	0.0686**	0.0373	0.0614*	0.0319	0.0137
	(0.0561)	(0.2484)	(0.0701)	(0.0342)	(0.0664)	(0.0319)	(0.0589)	(0.0298)
D˙Jambi	0.1153*	0.2029	0.1962**	0.0721*	0.0334	0.0145	0.0601	0.0631*
	(0.0660)	(0.2871)	(0.0833)	(0.0400)	(0.0786)	(0.0352)	(0.0682)	(0.0358)
D˙NTB	0.0628	-0.1858	0.0944	-0.0265	-0.0111	-0.0508	0.1020	-0.0436
	(0.0741)	(0.3295)	(0.0954)	(0.0472)	(0.0887)	(0.0424)	(0.0773)	(0.0406)
D˙NTT	0.1643**	-0.0116	0.2605***	0.0110	0.0837	-0.0225	0.2196***	0.0928**
	(0.0713)	(0.0377)	(0.0889)	(0.0434)	(0.0840)	(0.0386)	(0.0769)	(0.0402)
D˙Bappeda	0.0655	0.0035	0.0174	0.0263	0.0025	0.0121	-0.0436	-0.0372
	(0.0803)	(0.0427)	(0.0987)	(0.0431)	(0.9321)	(0.0436)	(0.0849)	(0.0423)
D˙Kehut	-0.0302	0.0560	-0.1231	-0.0007	-0.0333	0.0448	-0.1160*	-0.0206
	(0.0661)	(0.0372)	(0.0865)	(0.0431)	(0.0796)	(0.0383)	(0.0691)	(0.0360)
D˙LH	0.0197	0.0723*	0.0636	0.0914**	-0.0572	0.0554	0.0865	-0.0054
	(0.0715)	(0.0382)	(0.0890)	(0.0422)	(0.0861)	(0.0391)	(0.0765)	(0.0368)
Grad	0.2021	0.1672*	0.4015*	0.2086**	0.5096***	0.1411	0.1387	-0.0514
	(0.1673)	(0.0926)	(0.2096)	(0.1051)	(0.1943)	(0.0942)	(0.1770)	(0.0893)
Postgrad	0.7254	-0.0383	0.5732	0.0401	1.0491**	-0.1205	0.7351*	0.0353
	(0.4268)	(0.2278)	(0.5198)	(0.2587)	(0.5009)	(0.2374)	(0.4413)	(0.2146)
Observations	135	131	136	134	134	131	135	134
Prob $> \chi^2$	0.0638	0.0154	0.0034	0.0015	0.0065	0.0945	0.0334	0.0207
Pseudo-R^2	0.0604	0.0569	0.0982	0.0943	0.0877	0.0432	0.0766	0.0537

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Gedung LPEM FEB UI
Jl. Salemba Raya No. 4, Jakarta 10430
Phone : +62-21 3143177 ext. 621/623;
Fax : +62-21 3907235/31934310
Web : http://www.lpem.org/category/publikasi/workingppers/