Sustainable HRM through Improving the Measurement of Employee Work Engagement: Third-Person Rating Method

Feng Zeng Xu 1,2, Yun Zhang 1, Huixin Yang 1,* and Bob T. Wu 3

1 School of Management, Shandong University, 27 Shanda Nan Road, Jinan 250100, China; xfx@sdu.edu.cn (F.Z.X.); zhangyun8696@163.com (Y.Z.)
2 Center for Service Strategy and Service Management, Shandong University, 27 Shanda Nan Road, Jinan 250100, China
3 Allen W. and Carol M. Schmidthorst College of Business, Bowling Green State University, Bowling Green, OH 43403, USA; bwu@bgsu.edu

* Correspondence: yhx@sdu.edu.cn

Received: 9 July 2020; Accepted: 28 August 2020; Published: 31 August 2020

Abstract: The purpose of the paper is to present the survey findings of two alternative methods (self-rating (SR) and third-person rating (TPR)) of measuring employee work engagement (EWE). The potential impacts of gender, job tenure, position, and work condition on TPR vs. SR were also investigated. A sample of 649 of hotel service workers, supervisors, and managers in China participated in the study. An accurate measure of employee work engagement serves as a leading indicator of turnover intention and an early diagnostic tool for sustainable human resource management. Despite its popularity as a work engagement measure, SR method has many limitations. This research attempted to demonstrate that TPR is a viable and better alternative measure of EWE. The results indicated that TPR does possess desirable measurement characteristics, such as convergent validity, nomological validity, and structure invariant. TPR also provides a more conservative, and perhaps more accurate as well, measure of EWE. The difference in mean EWE scores as measured by SR vs. TPR was found to be affected by the specific dimension under study, with the least observable absorption dimension the most affected. The difference was also found to be significantly higher for males than for females, bigger as an employee’s position moves higher, and larger as the length of job tenure increases. Additionally, the difference in satisfaction–EWE correlations, as measured by TPR vs. SR, were much higher when the work conditions were poor. For practitioners, the importance of this study lies in the fact that TPR, as a conservative measure of EWE, can play an important role in detecting early signs of employee troubles sooner and lead management to take timely actions, making human resource management more sustainable. For academics, the results that SR and TPR of EWE generally result in similar pattern of findings offer strong encouragement to build future research on EWE through the TPR method.

Keywords: employee work engagement; sustainable human resource management; third-person rating; self-rating; hospitality

1. Introduction

Sustainability, as one of the fastest growing research areas, has received considerable attention from both practitioners and researchers [1,2]. However, there is no universal definition of sustainability. Sustainability can be simply referred to as something that is capable of being continued at a certain level for a foreseeable future. Economy, society, and the environment are the three pillars of sustainability. For example, sustainability can be applied to study how humankind avoids the depletion of natural
resources, the decrease of society’s quality of life, and the recession or depression of economy. All three pillars are interrelated. The decisions made for one pillar will have consequences on other pillars as well. As evidence of the importance and urgency of sustainability issues facing the whole world, the world leaders agreed in 2015 on seventeen Sustainable Development Goals (SDGs) [3]. The goal #8 for UN’s 2030 agenda called for all nations to promote inclusive and sustainable economic growth, employment, and decent work for all [3]. Being on the supply side of the economy, business organizations are on the frontline for the attainment of goal #8—sustainable economic growth and the creation of inclusive decent work for all. The attainment, in turn, depends very much on sustainable human resource management (SHRM). Therefore, one important individual and collective human sustainability issue is sustainable human resources management. SHRM is described as the “adoption of HRM strategies and practices that enable the achievement of financial, social and ecological goals, with an impact inside and outside of the organization and over a long-term time horizon while controlling for unintended side effects and negative feedback” [4] (p. 90). The long-term satisfaction and retention of employees is an important domain of SHRM. Research has generally showed a positive linkage between employee work engagement (EWE), work satisfaction, and stay intention [5,6]. As a result, an important SHRM issue is to increase EWE, reduce turnover rates, and achieve sustainable long-term employee retention. Such relationships have led to a great interest by both academics and practitioners in monitoring the level of EWE. EWE was conceptualized as a positive state of employee motivation, through which people employ and express themselves physically, cognitively, and emotionally during role performances [7], or as a positive antithesis of burnout [8]. The antecedents of EWE have been conceptualized as job characteristics, perceived organizational and supervisory supports, procedural and distributive justices, and rewards and recognitions [9].

The general level of employee engagement varies from industry to industry. Employment in the tourism and hospitality industry is characterized by a heavy workload and stress, working unsocial hours, less respect, and low rates of pay [10], which in turn lead to instability in contracts and high levels of turnover [11]. Thus, human resource management in the tourism and hospitality industry is particularly challenging and very much relevant to the study of sustainability for practitioners and academics alike. Previous studies have made great advancements on the development of sustainable tourism [12–14]. However, extant research has been conducted primarily through ecological and social perspectives of tourism [15–17], while individual and collective human sustainability of the hospitality industry has received little attention [18]. The present study fills the gap by studying the role EWE measurement plays in SHRM under the hospitality setting.

EWE measures can serve as a leading indicator of turnover intention and an early diagnostic tool of sustainability. For example, according to the latest engagement statistics, only 15% of workers are engaged in their work and 81% of workers consider leaving their job [19]. Similarly, the Gallup’s State of the Global Workplace survey in 2019 showed that the percentage of employees who are engaged is 31% in the United States and Canada, while only 6% in East Asia [20]. The low level of EWE, particularly in Asia, calls into question the long-term sustainability of business survival and growth. It also points to the needs to examine the validity of the EWE measures and the factors influencing such measurement.

Most of these engagement surveys used self-rating (SR) as the measurement instrument. However, there are conflicting reports on the adequacy of self-rating-based surveys. Research has shown that self-ratings of performance tend to be more lenient than supervisory ratings [21]. Consequently, there are calls to employ multi-item measures using other methods, such as supervisors and others, to measure key management constructs [22]. Despite such calls for the employment of other methods and the potential limitation of the self-rating method, scant research has investigated the adequacy of self-rating in the context of EWE measurement. In her study on engagement, Wilson [23] explicitly recognized this weakness by indicating “there are notable limitations to this study. It is understood that biases may be inherent in the self-reported information” (p. 10). Nevertheless, she provided no further discussions on such bias. Nor did she offer methods to correct the limitation. Given the popularity of
SR measure and its potential limitations, more research is needed to explore alternative methods to measure the construct of EWE.

One potential problem of using SR to measure EWE is the social desirability bias. EWE measures involve sensitive inquiries about potentially self-incriminating information. Therefore, although employees themselves have the best knowledge of their own level of engagement, one concern is the possibility that employees over-report the extent to which they engage in works, either to bolster self-esteem or for self-serving purposes. One technique to overcome such bias may be the use of the third-person rating (TPR) technique [24]. The TPR technique is used to elicit true opinions held by respondents that might be perceived as reflecting negatively upon the individual. However, not much is known about TPR as a measure of EWE. For TPR to be a viable alternative measure, following questions need to be answered: Is there a significant correlation between the results of TPR and SR? Does TPR overcome the over-reporting bias of SR? To what extent does the EWE obtained from SR correlate to the EWE assessed through TPR?

Having an accurate diagnostic device is the essential first step to contain and cure regular diseases or a pandemic to attain a sustainable national health. False negative (infected by a disease without being detected due to fail diagnostic device) prevents timely medical treatment from being administered. By the same token, applying the right and accurate EWE measurement method is also the essential first step to send management the timely warning sign for early intervention. Inherent optimistic measures such as the SR method suffers the false negative error by inflating the level of EWE. Being conservative, TPR method offers management a timely and accurate reading of the health of employees’ motivational states, thereby, warning management for timely intervention to achieve employee satisfaction and sustainability. Through the use of both SR and TPR methods in a questionnaire survey over 649 hotel employees, we empirically investigated a number of important issues regarding the degree to which using SR vs. TPR makes a difference in EWE research. For example, do SR and TPR report different levels of EWE? Furthermore, EWE is a multidimensional construct [25], some dimensions are more objectively observable than others (e.g., vigor vs. absorption). Do SR and TPR produce more similar level in an objectively observable EWE dimension than in dimension that is hard to observe directly and objectively? Also, are there certain variables expected to moderate the results produced by SR vs. TPR in EWE measurement? If so, what are those variables? How are the results affected? Thus, the primary objectives of the present study were: (1) to investigate the potential bias of SR EWE measure by comparing SR to TPR; (2) to compare the dimensionality of EWE measures under SR vs. TPR methods; and (3) to explore the factors that may regulate the degree and direction of results obtained from SR vs. TPR.

2. Literature Review

The relationship between sustainability and sustainable human resource management rests on the assumption that the practices and strategies of sustainable human resource management will have a significant impact on natural and social resources, as well as on the management conditions that affect business’ long-term growth [26]. In the perspective of the stakeholders’ theory, it is expected that the human resource managers will take into consideration the interests of its stakeholders when making business decisions. Specifically, the policies and actions will focus on contributing to long-term survival and corporate success, the workplace well-being, and benefiting society [27]. The stakeholder orientation of SHRM also favors the adoption of processes fostering the perception of procedural and distributive justice in the organization, employee satisfaction and retention, etc. [28]. Such perceptions are some of the antecedents for EWE, with consequences of job satisfaction and turnover reduction [9]. EWE clearly plays a key mediator role between human resource management’s policies and actions and the goal of pursuing sustainable employee wellbeing.
2.1. Work Engagement

Engagement has been defined in many different ways. It has been defined as “emotional and intellectual commitment to the organization” or as “the amount of discretionary effort exhibited by employees in their jobs” [9] (p. 601). Given the hospitality context of the current study, employee work engagement is defined as the extent to which employees are involved with, enthusiastic about, and committed to their work [29]. It was first conceptualized by Kahn [30] to measure how “people can use varying degrees of their selves physically, cognitively, and emotionally in the roles they perform” (p. 692). Implicit in Kahn’s conceptualization is a three-dimensional factor structure. A highly engaged employee is characterized as an employee who has a high level of energy (vigor), identifies strongly with his/her work (dedication), and is fully immersed in his/her job (absorption) [29]. Most of the engagement measurement adopt the three-dimensional conceptualization, such as Schaufeli and Bakker [25]. Despite the popularity of the three-dimensional conceptualization, there are research studies that failed to replicate the three-factor structure of work engagement [31].

Research on work engagement generally falls into three areas. The first area is the debate of the conceptual uniqueness of work engagement from other related concepts and the measurement of the concepts [32,33]. Some scholars described work engagement as “old wine in a new bottle” because of the conceptual overlap with the traditional workplace-related attitude [9]. For example, González-Romá et al. [34] claimed that the three dimensions of work engagement (energy, involvement, and efficacy) are the opposite pole of the three corresponding aspects of burnout (exhaustion, cynicism, and inefficacy); Joseph et al. [35] indicated that several statements of the measure of work engagement (e.g., Dedication 2 on the original UWES) overlapped with the Brayfield and Rothe’s [36] (p. 311) statements of job satisfaction. While some other scholars insisted that work engagement is related to but distinct from these concepts. Schaufeli and Bakker [25] explained that the two concepts of work engagement and burnout cannot be perfectly negatively correlated, so that they are two distinct concepts and should be assessed independently. Christian et al. [37] provided the evidence that employee engagement and job satisfaction are two distinct concepts through a meta-analytic work. Moreover, work engagement is not merely an attitude but involves cognitions, emotions, and behaviors, which distinguish themselves from the concepts of commitment, job involvement, and organizational citizenship behavior [9]. Subsequent literature has shown that work engagement in academic literature has been commonly accepted as a distinct and unique construct [9]. The related constructs are explored as the consequences associated with work engagement [9].

The other two areas of EWE research pay attention to the predictors leading to EWE and the consequences of EWE [38–42]. Regarding the predictors of EWE, there are three streams of research. The first stream focuses on the conditions of job which leads to work engagement. Job conditions that are more psychologically available and offer employees more psychological meaningfulness and safety will lead to higher employee engagement [30]. The second stream comes from the burnout literature. The following six aspects of work life were found to lead to engagement or disengagement: workload, control, rewards, recognition, community and social support, perceived fairness, and values [8]. The third stream focuses on the impacts of employee personality on EWE. Scholars have indicated that personality characteristics (e.g., positive affectivity, motivation, and psychological capital) exert a significant and positive impact on EWE [33,40,41]. The factors identified in the first two streams are antecedent predictors, while the third stream is trait predictors but not antecedents. As for the consequences of EWE, the social exchange theory offers some insights. According to SET, people engage in social exchange with others to maximize benefits and minimize costs. Reciprocity is a key component in such well calculated behaviors. In the context of business organizations, when employees receive support from their organization or supervisors, they feel it is in their best interests to reciprocate the organization with greater levels of engagement [9]. More engagement brings more favorable supports from their supervisors and organization, leading to a virtuous cycle. Thus, employees are more likely to report more positive attitudes and intentions toward the organization when highly engaged [8].
Based on the first two streams of research, Saks [9] proposed the antecedents→EWE→consequences model, as described in the top portion of Figure 1.

![Figure 1](image_url)

**Figure 1.** The conceptual level and measurement level framework of employee work engagement (EWE).

For work engagement to serve as a leading indicator of turnover intention and an early diagnostic tool of sustainability, its measurement has to be valid, thus this study focused on the first area of engagement research.

2.2. Third-Person Technique

Most researchers agree that method variance is a potential problem in behavioral research [43,44]. Method variance refers to variance that is attributable to the measurement method rather than to the constructs the measures represent [43]. Bias due to the method used is a major source of measurement error, which has both a systematic component and a random component. Method bias represents the systematic measurement error. The impact of method bias is to either inflate or deflate observed relationships between constructs. Such distortion can have a serious confounding influence on empirical results, yielding potentially misleading conclusions. The common sources of method bias include the context in which the measures are obtained, social desirability, leniency effect, etc. Social desirability refers to the tendency of subjects to respond favorably to items more as a result of their social acceptability than their true feelings [45]. Leniency refers to tendency to rate those whom they know well, or are ego involved, higher than they should [46]. The amount of variance attributable to method biases varied considerably by the method used and by the type of construct being investigated.

Fully engaging in one’s work is deemed to be highly desired and expected by the society. EWE as a construct has the social desirability characteristics and its measurement can be highly inflated, particularly if the method of measurement is self-rating. There are many techniques for controlling common method biases. The two primary ways are through the design of the study’s procedures and/or statistical controls [43]. An often overlooked remedy is to change the context in which the questions are framed. Thus, this study focused on one frame-changing method: the third-person rating technique. In certain circumstances, it is quite unlikely to obtain accurate information about what people think and feel by direct questioning [47]. Some social norms, self-interests, or negative consequences may prevent respondents from freely expressing their true feelings or opinions [48]. Under such circumstances, the third-person technique (also being referred to as projective questioning [49], can be deployed to solicit the true sentiments. Instead of questioning them directly, the subjects may be asked to respond
indirectly, discussing about others’ feelings, attitudes, and opinions. In talking about a third party, the subjects project their covert feelings to the third party [50] (p. 136).

2.3. Third-Person Effect

The use of third-person technique may also trigger the self–others discrepancy phenomenon, also called the third-person effect (TPE) [51]. It is hypothesized that people exposed to persuasive mass media messages will perceive these messages to have more influence on other people than on themselves [52]. To gauge the TPE, researchers often ask subjects pairs of questions. For example, subjects may first be asked how much they think the message has affected their opinions. Then they are asked how much they think the same message would affect the opinion of others who see the message. If the response to the second question is greater than the response to the first question, a TPE is assumed. The self–others discrepancy can be due to overestimations of impact on perceived others, underestimations of impact on oneself, or both. Research more often supports overestimation of the impact on perceived others’ view [53–56]. The overestimation explanation views others as “vulnerable” to external influence [57]. Although there are several psychological theories hypothesized to be the underlying mechanism for TPE, one that is particularly relevant in the current context is the concept of biased optimism [58]. Biased optimism refers to the idea that “...human tendency to see the world through optimistic or self-serving lenses... will estimate greater media effects on others...for messages with harmful outcomes, but no difference in effect for beneficial message.” [58] (p. 58). In the context of EWE, employees may perceive fellow workers to be more vulnerable to various external influences to fully engage in office works than they do.

3. Hypotheses Development

Dedication to one’s work is both socially desirable and potentially affecting an employee’s chance of getting a reward/promotion; therefore, respondents, when asked to self-rate themselves, will evaluate themselves highly engaged to meet the social desirability norms or for self-serving purpose (self-enhancement motivation). Research indicated that respondents have the tendency to inflate the rating of their own performance compared to ratings from supervisors [59,60]. Several theoretical explanations have been offered to explain why the leniency of self-rating exists. Alicke et al. [61] referred to the tendency to evaluate oneself more positively than others as the “better-than-average effect”. Other researchers indicated that self-leniency is a way of enhancing one’s self-image (self-esteem) [21]. Regardless of the reason behind the leniency, the existing research does suggest a systematic, positive bias effect of self-rating method. In other words, if $x_t$ represents the one’s true engagement level, and $y_s$ is the reported score from self-rating, then, $y_s = x_t + c_s + e_s$, where $c_s$ is a positive constant, representing the positive bias, and $e_s$ the random error term.

On the other hand, when subjects are asked to rate the engagement level for “others”, they are either free to project their own true feeling without the positive bias or the TPE may push an overestimation of others of being vulnerable, of not working hard, a negative bias. That is, if $x_t$ represents others’ (third-person) true engagement level, and $y_t$ reported score from third-person’s rating, then, $y_t = x_t + c_t + e_t$, where, $c_t$ is a negative constant, representing the negative bias, and $e_t$ is the random error term. Collectively, the true mean of all individuals in a group should be the same as the true mean of all others in the same group, therefore, the mean of $x_t$ should be the same as the mean of $x_t$, and the mean of $y_s$ should be higher than the mean of $y_t$ by the sum of $c_s - c_t$. $Y_t$ is basically a linear transformation of $y_s$ with a constant term and an error term. There should be a very high correlation between $y_s$ and $y_t$. In their study of organizational citizenship behavior (OCB), Khalid and Ali [62] reported a correlation coefficient of 0.35 between overall OCB measured by self-ratings vs. by superior ratings. Similarly, in a meta-analysis of counterproductive work behavior (CWB), Berry et al. [47] reported that the average corrected self-other (coworkers or supervisors) CWB correlation was 0.38. Given that $y_s$ and $y_t$ are two separate ratings from the same source measuring the same construct, it is expected that common method bias will inflate the correlation between the results from the two
methods. Thus, this correlation should also be higher than the correlation between ratings from two sources, such as self-rating vs. supervisory rating. Therefore,

**Hypothesis 1 (H1).** *There is a significant positive relationship between employee work engagement measured by self-rating and employee work engagement measured by third-person rating.*

Given that \(x_s = x_t\), and \(c_t\) is a negative value, \(y_s - y_t = c + e\) (where \(c = c_s - c_t\), and \(e = e_s - e_t\)) will be larger than zero; that is \(y_s\) is larger than \(y_t\). Furthermore, the more objectively observable the dimension of work engagement, the less likely the rating of such dimension will be subject to a perceptual distortion [24]. In general, outer behavior is easier to observe than the inner mind state. Among the three engagement dimensions (vigor, dedication, and absorption), vigor refers to an employee’s level of energy and is probably the most objectively observable through an employee’s outer active behavior. Dedication is the degree of identification with one’s work and is manifested by an employee’s outspoken endorsement and support of his/her job. Absorption refers to how much an employee is fully immersed in his/her job, the least observable inner mind state inferred from an employee’s outer concentration and focus on works. A high level of energy (vigor) exhibited by an employee is much easier to be observed by everyone than how much an employee immerses in his/her job. Therefore:

**Hypothesis 2 (H2a).** *Employee work engagement measured by self-rating is significantly higher than employee work engagement measured by third-person rating.*

**Hypothesis 2 (H2b).** *The differences in employee work engagement measured by self-rating vs. by third-person rating is largest for the dedication dimension and smallest for the absorption dimension.*

For a method to be a viable alternative measure, in addition to convergent validity, the measure has to exhibit a nomological validity by correlating other constructs in a theoretically expected direction. One of the antecedent variables for work engagement is an employee’s satisfaction with the work environment, a manifestation of organizational supports. The more satisfied an employee is with the work environment, the more engaged an employee will be with his/her works [63]. Therefore, EWE, whether measured by SR or by TPR, should both correlate positively with work environment satisfaction. Furthermore, according to biased optimism hypothesis, the third-person effect is most profound when a person is exposed to harmful messages [58]. In the context of EWE measurement, when an employee is exposed to an unfavorable work environment, he/she will perceive his/her fellow workers will be affected more negatively by the poor work environments than himself/herself and will show a lack of engagement in the job. Thus, the differences in EWE measured by SR versus by TPR will be much larger when work environments are perceived to be poor than when environments are perceived to be good. Therefore:

**Hypothesis 3 (H3a).** *Employee work engagement is positively related to work environment satisfaction for both self-rating and third-person rating.*

**Hypothesis 3 (H3b).** *The differences in employee work engagement measured by self-rating vs. by third-person rating is significantly larger under the poor work environment condition than under the good work environment condition.*

The gender of a supervisor or manager has been reported to bias the ratings of work performance [64]. The positive bias hypothesized in H2a may be moderated by respondents’ gender. After an extensive review of research findings, Dipboye and Flanagan [65] asserted that the positive bias from SR as compared to supervisory may apply only to male respondents because most of the past research had men as subjects. Subsequently, Waldman, as reported by Shore and Thornton [66] (p. 116), found that SR was lower (rather than higher) than supervisory ratings when the instrument was
administered to women-only subjects. Shore and Thornton’s own research [66] found no significant gender differences in both SR and supervisory ratings. Despite the inconsistent results, the majority of the research findings indicate that male subjects tend to rate themselves higher than female subjects [64,67,68]. Therefore, males can be expected to show a tendency to inflate self-ratings compared to that of females. That is, \( c_s \) is larger for males than for females.

**Hypothesis 4 (H4a).** Employee work engagement measured by self-rating is significantly higher for males than for females.

Furthermore, the socialization of gender roles suggests that women are socialized to behave in a more compassionate, nurturing, and cooperative manner, whereas males are generally socialized to be competitive and independent [69,70]. Shore and Thornton [66] (p. 118) asserted “The majority of studies investigating raters’ genders have shown a trend for women to rate others higher than do men”. Consequently, it can be expected that the negative bias associated with TPE as hypothesized in H2a will diminish for females more than for males [54]. That is, the negative \( c_t \) for males is larger than for females. Therefore:

**Hypothesis 4 (H4b).** Employee work engagement measured by third-person rating is significantly lower for males than for females.

The combination of self-boasting (larger \( c_s \)) and being more critical of others (larger negative \( c_t \)) by males as compared to females will result in more profound differences in scores between SR and TPR for males than for females. Therefore:

**Hypothesis 4 (H4c).** The differences in employee work engagement measured by self-rating vs. by third-person rating is significantly larger for males than for females.

Employees started off energized and engaged with their jobs but tended to drop off a bit after six months, and substantially after ten years of tenure [71]. Similarly, Attridge [72] reported a negative relationship between the number of years an employee worked and the degree of engagement. Such relationship may color an employee’s judgment of the level of engagement of self or others. The lower level of work engagement makes employees with longer tenure less secure and feel a stronger need to be more critical of others in order to protect themselves than do those with short tenure. Gunther and Mundy [58] (p. 58) reported that “people will estimate greater media effects on others than on themselves for messages with harmful outcomes, but no difference in effect for beneficial messages”. Such observation may be extended to hypothesize that when employees are in low engagement situations (potential harmful outcomes) due to longer tenure, they will expect that others are at a much worse engagement level than they are. Thus, EWE score measured by SR will be much higher than EWE score obtained through TPR. On the other hand, for those with higher engagement level (beneficial outcomes) due to short tenure, EWE score measured by SR will be higher than, but very similar to, the score assessed by TPR.

However, the “depletion” effect of service years on engagement was not observed in Wilson’s study [23]. Coffman and Gonzaler-Molina’s [71] assertion will be tested again in this study. Thus:

**Hypothesis 5 (H5a).** Employee work engagement is negatively related to employee’s years of employment.

**Hypothesis 5 (H5b).** The difference in employee work engagement measured by self-rating vs. by third-person rating is significantly higher for employees with longer tenure than those with short tenure.

Kerfoot [73] suggested that superiors set behavioral examples for their subordinates to follow. If a leader is not engaged on the job, it will provide excuses for his/her subordinates not to engage.
Furthermore, Saks [9] reported that perceived organizational support exerted a strong positive impact on employee engagement. A lowly engaged leader is not likely to provide strong support to the subordinates, and thus is not likely to have engaged subordinates. Finally, employees with leadership positions are paid much higher than workers occupying nonleadership positions and are treated with respect by their subordinators. Thus, employees with leadership positions are more likely to perceive themselves to be highly engaged in works (beneficial outcomes). Given Gunther and Mundy’s [58] assertion that little third-person effect is observed for beneficial messages, it can be expected that those with leadership positions will perceive similar engagement levels for themselves and their fellow leaders, EWE score measured by SR will be higher than but very similar to the score assessed by TPR. Therefore:

**Hypothesis 6 (H6a).** Employee work engagement is higher for managers and supervisors than for the non-manager/supervisors.

**Hypothesis 6 (H6b).** The difference in employee work engagement measured by self-rating vs. by third-person rating is significantly lower for employees with leadership positions than those without leadership position.

Figure 1 illustrates the relationships between antecedents–EWE–consequences (A–E–C). It depicts the relationship between A–E–C both at the unobserved construct level (top portion above the dotted line) and at the observable measurement level (bottom portion below the dotted line). The observable measurement scores are influenced by the underlying constructs (antecedent variables, EWE, consequence variables), measurement methods (SR vs. TPR), demographics (gender, work tenure, and job position), and random errors. Additionally, the demographics and one predictor (satisfaction with the work condition) also serve as moderators regulating the effect of measurement methods on EWE. Since researchers use the observable scores and correlations among scores to infer the level of unobservable constructs and the relationships between constructs, it is important that the measurement scores obtained truthfully reflect the underlying constructs and relationships. Given that this study focused on investigating bias due to alternative methods of measuring EWE and how a selected variable might moderate such bias, the majority of hypotheses deal with the levels of, and relationships between, variables at measurement level (bottom portion, Figure 1; H1, H2a,b, H3b, H4a,b,c, H5b, H6b). The rest of hypotheses are at conceptual/construct level and not method specific (H3a, H5a, H6a).

4. Method

4.1. Research Design and Setting

The research design was a field study using survey methodology. The study was carried out with a sample of employees (front line employees, supervisory staff, and managers) from hotels in a large metropolitan city in northeastern China. The survey questionnaire was designed to measure the participants’ perception about their own levels of engagement (SR), their satisfaction with the work environment, organizational support, supervisory support, rewards and recognition, and their agreement with various descriptions of the job characteristics. They were also asked to indicate the levels of engagement of their fellow workers (TPR). Demographic questions (i.e., gender, age, education, marital status, length of working at company, and job position) were asked at the end of the questionnaire. Since authors are fluent in both Chinese and English, parallel translation was used to translate English questions into the Chinese version. Parallel translation has been advocated as a preferred method of achieving equivalence in meaning in international studies [74].

4.2. Sample and Procedure

The study was carried out in a large metropolitan city in northeastern China. Using a convenient sampling method characterized by easy accessibility, geographical proximity, and availability at a given time [75], the research team contacted senior executives of ten hotels, and eventually, seven hotels
agreed to participate in the survey. These hotels belong to seven different brands and have been set up from seven years to twenty-seven years.

The research team distributed questionnaires to employees face-to-face with the support of senior executives of these seven hotels. The participants were told that the investigation was anonymous and the questionnaires were self-administered. Then, the participants were left to complete the questionnaire independently. Of the 800 distributed questionnaires, 722 were returned. After screening and excluding the invalid questionnaires, 649 were retained for data analysis.

4.3. Variable Measures

Utrecht Work Engagement Scale (UWES) is the most often used instrument internationally [76]. It has been validated in countries of various cultural backgrounds, including China [76]. The UWES includes three dimensions: vigor, dedication, and absorption. Studies confirmed that the fit of the three-factor structure to the data was superior to that of any other alternative factor structures [76]. There are three versions of the UWES scale: 17 items, 15 items, and 9 items [25] (p. 14). The original long version of UWES includes 6 vigor items, 5 dedication items, and 6 absorption items. Subsequent psychometric analyses uncovered two weak items, so that some studies only used a 15-item version. In order to shorten the scale further, the top three items with highest loadings on each dimension were selected by some scholars, which resulted in a 9-item scale. The shorter version was ideal for encouraging participation in the study. However, the three dimensions of 9-item scale have been found to be strongly related to each other, resulting in some studies failing to find the three-factor structure [76]. Consequently, a slightly longer than 9 items version was needed for the present study. We found a version improved over the original 17 items by Salanova et al. [77]. We adopted their version. After a pretest, the two items from the vigor scale (items 1 and 3) were deleted because the items referred to “energy” while working and liking for work [77] (p. 1222). Item 5 and 6 from the absorption scale were deleted because of the similarity of content with item 4. Item 2 from the dedication dimension was also excluded for the repeating semantics. Finally, the remaining 12 items in the original version were used, with four items in each of the three dimensions. (see Table 2 for the list of items).

Following Lee and Chang [78], two environment condition satisfactions were measured: work environment condition and living environment condition. Work condition satisfaction was measured by five items with a scale from very satisfied (5) to very dissatisfied (1). Living condition satisfaction was measured by four items using the same five-point scale as work condition.

5. Results

5.1. Descriptive Statistics

A total of 649 employees from seven hotels participated in the study. As shown in Table 1, the respondents comprised of 45.9% males and 54.1% females. Most of the respondents were aged 21–30 years (56.9%), followed by those aged 31–40 years (25.0%). In terms of education, 56.8% of the respondents had an educational level of high school or below, and 32.1% and 11.1%, had completed education in vocational schools and universities, respectively. Among the respondents, 43.8% were married and 56.2% were single/divorced. The distribution of the length of time working at the hotels was as follows: 19.5% <1 year, 18.5% 1–2 years, 24.1% 3–5 years, 14.6% 6–9 years, and 23.3% ≥10 years. For job position, there were 33.8% frontline service staffs, 12.2% frontline operation staffs (cooks, dishwashers, etc.), 10.5% maintenance personnel, 22.9% low level supervisors, 11.9% mid-level managers, 3.9% upper-level managers, and 4.8% miscellaneous job positions.
### Table 1. Sample characteristics.

| Characteristics                  | Frequency (n) | Percentage (%) |
|----------------------------------|---------------|----------------|
| **Gender**                       |               |                |
| Male                             | 298           | 45.9           |
| Female                           | 351           | 54.1           |
| **Age (years)**                  |               |                |
| ≤20                              | 63            | 9.7            |
| 21–30                            | 368           | 56.9           |
| 31–40                            | 162           | 25.0           |
| 41–50                            | 50            | 7.7            |
| >50                              | 4             | 0.6            |
| **Education**                    |               |                |
| Middle school or lower           | 74            | 11.6           |
| High school                      | 289           | 45.2           |
| College                          | 205           | 32.1           |
| Undergraduate or higher          | 71            | 11.1           |
| **Marital Status**               |               |                |
| Married                          | 284           | 43.8           |
| Single/divorced                  | 364           | 56.2           |
| **Length of Working at Company (years)** | | |
| <1 year                          | 126           | 19.5           |
| 1–2                              | 119           | 18.5           |
| 3–5                              | 155           | 24.1           |
| 6–9                              | 94            | 14.6           |
| ≥10                              | 150           | 23.3           |
| **Job Position**                 |               |                |
| Frontline service staffs         | 217           | 33.8           |
| Frontline operation staffs       | 78            | 12.2           |
| Maintenance personnel            | 67            | 10.5           |
| Low level supervisors            | 147           | 22.9           |
| Mid-level managers               | 76            | 11.9           |
| Upper-level managers             | 25            | 3.9            |
| Miscellaneous job positions      | 31            | 4.8            |

5.2. Reliability, Validity, and Measurement Model Fit

Although UWES has been validated across a variety of countries and industries [25], analyses were performed to validate the 12 item UWES scale to ensure its adequacy for the current application (hospitality industry in China). Using Amos 23.0, confirmatory factor analysis (CFA) was performed on the data collected. As shown in Table 2, for self-rating (SR), all items had standardized factor loadings larger than the suggested 0.5 level [79] (p. 686). The indices associated with the revised model met the respective criteria (NFI = 0.95; CFI = 0.96; IFI = 0.96; TLI = 0.94; RMSEA = 0.07; SRMR = 0.03). The Cronbach’s alphas for the items associated with the three dimensions were 0.73, 0.88, and 0.78, all above the threshold value of 0.7 were deemed sufficient for basic research [79].

For the third-person rating (TPR), the CFA results showed that the standardized factor loadings of all items were higher than 0.5. Two items with absorption were highly correlated according the value of modification index, and the first was dropped from further analyses. The indices associated with the revised model were NFI = 0.96; CFI = 0.97; IFI = 0.97; TLI = 0.96; RMSEA = 0.08; SRMR = 0.04. The final list of items for each dimension of TPR is shown in Table 2. The Cronbach’s alphas for the three dimensions were 0.85, 0.92, and 0.84 (Table 3). Additionally, the results of the correlations between the three dimensions from the same method were also quite comparable to the ranges of correlation values of UWES-9 reported in Schaufeli and Bakker’s study [25]. For example, for dedication–absorption: UWES-9 reported a range of 0.52 to 0.84, while for the present study, it was 0.75 for SR and 0.76 for TPR (Table 3). Therefore, the selected UWES scales for both SR and TPR in this study were valid measures of work engagement.
Table 2. Revised Utrecht Work Engagement Scale (UWES) scale for self-rating (SR) vs. third-person rating (TPR).

| Items                                                                 | SFL      | Mean          | S.D.       |
|----------------------------------------------------------------------|----------|---------------|------------|
|                                                                      | SR       | TPR           | SR         | TPR       |
| Vigor:                                                              |          |               |            |           |
| 1. At my job, I feel strong and vigorous.                           | 4.22     | 3.76          | 0.65       | 0.79      |
| 2. I can continue working for very long periods at a time.           |          |               |            |           |
| 3. At my job, I am very resilient, mentally.                        |          |               |            |           |
| 4. At my work I always persevere, even when things do not go well.  |          |               |            |           |
| Dedication:                                                         |          |               |            |           |
| 1. I am proud of the work that I do.                                |          |               |            |           |
| 2. To me, my job is challenging.                                    |          |               |            |           |
| 3. My job inspires me.                                               |          |               |            |           |
| 4. I find the work that I do full of meaning and purpose            |          |               |            |           |
| Absorption:                                                         |          |               |            |           |
| 1. I am immersed in my work.                                       |          |               |            |           |
| 2. When I am working, I forget everything else around me.           |          |               |            |           |
| 3. Time flies when I’m working                                      |          |               |            |           |
| 4. I feel happy when I am working intensely                         |          |               |            |           |
| Indices:                                                            | 0.95     | 0.96          |            |           |
| NFI                                                                  | 0.96     |               |            |           |
| CFI                                                                  | 0.96     |               |            |           |
| IFI                                                                  | 0.96     |               |            |           |
| TLI                                                                  | 0.94     |               |            |           |
| RMSEA                                                                | 0.07     |               |            |           |
| SRMR                                                                | 0.04     |               |            |           |

Notes: SR = self-rating; TPR = third-person rating; SFL = standardized factor loading.

Table 3. Correlation coefficient between self-rating vs. third-person rating.

|                  | 1  | 2  | 3  | 4  | 5  | 6  | 7  | 8  |
|------------------|----|----|----|----|----|----|----|----|
| 1. (SR) Overall Engagement | -  |    |    |    |    |    |    |    |
| 2. (SR) Vigor     | 0.83 | (0.73) |    |    |    |    |    |    |
| 3. (SR) Dedication | 0.90 | 0.59 | (0.88) |    |    |    |    |    |
| 4. (SR) Absorption | 0.91 | 0.67 | 0.75 | (0.78) |    |    |    |    |
| 5. (TPR) Overall Engagement | 0.67 | 0.52 | 0.63 | 0.62 | -  |    |    |    |
| 6. (TPR) Vigor    | 0.58 | 0.52 | 0.50 | 0.52 | 0.91 | (0.85) |    |    |
| 7. (TPR) Dedication | 0.65 | 0.45 | 0.68 | 0.56 | 0.92 | 0.74 | (0.92) |    |
| 8. (TPR) Absorption | 0.61 | 0.46 | 0.54 | 0.61 | 0.91 | 0.76 | 0.76 | (0.84) |

Notes: SR = self-rating; TPR = third-person rating; The values of Cronbach's alpha are in the diagonal.

Although the composition of items for SR is slightly different from that of TPR, this is not uncommon. In studying workplace deviance behavior, Stewart et al. [22] compared self-report measures with non-self-report (fellow workers, supervisors) measures. They reported that, “… the factor structure may become more refined when reporting others’ behaviors in comparison to self-reporting behavior …” (p. 213). They used fundamental attribution bias to help explain the difference “… when one examines others’ behaviors, different cognitive activities may be performed to understand events … in comparison to understanding … one’s own behavior” (p. 213).

Before performing tests on the hypotheses, the item scores of each dimension were averaged to serve as a summated measure of that dimension. An overall engagement score was also generated by averaging the three dimensions score as suggested by Schaufeli and Bakker [25] (p. 33).

Table 3 presents the Pearson correlation coefficients between the three dimensions (vigor, dedication, and absorption) and overall engagement for SR and TPR. Both methods showed a high
degree of internal consistency. Individual dimension scores were highly correlated with the overall score, ranging from 0.83 to 0.91 for SR and from 0.91 to 0.92 for TPR. Furthermore, the correlations between ratings by two different methods (SR vs. TPR) on the same trait (dimension) were mostly high, particularly at the overall engagement level (overall engagement: $r = 0.67$; vigor: $r = 0.52$; dedication: $r = 0.68$; absorption: $r = 0.61$; all correlations: $p < 0.00$), a sign of convergent validity.

5.3. Hypotheses Testing

H1 proposed that there is a high positive correlation between self-rating and third-person rating of employee work engagement. For comparison purposes, Khalid and Ali [62] reported a correlation coefficient of 0.35 between OCB measured by self-ratings vs. by superior ratings, and the correlations between different methods on various dimensions of OCB were all below 0.35. Berry et al. [47] reported that the average corrected self-other (coworkers or supervisors) CWB correlation was 0.38. Therefore, H1 was supported as there was a significantly positive correlation (>0.38) between SR and TPR of employee engagement.

Table 4 shows the means for the three dimensions and the overall engagement score for both methods. Ratings from SR were consistently and significantly higher than ratings from TPR across all three dimensions and the engagement measure (paired t-tests, $p < 0.00$). The overall engagement level by SR was 4.13 (out of a maximum of 5), while the score is 3.65 for TPR, a clear “better-than-average” positive bias [61]. Therefore, H2a was supported as EWE measured by SR was significantly higher than EWE by TPR. Significant differences between SR and TPR were also found for each of the three dimensions. As expected, the differences between SR and TPR were the smallest for vigor dimension (0.46), followed by dedication (0.47) and absorption dimensions (0.53). H2b was supported.

Table 4. Mean ratings for vigor, dedication, absorption, and overall engagement. Self-rating vs. third-person rating.

| Overall Engagement | Vigor | Dedication | Absorption |
|--------------------|-------|------------|------------|
| SR                 | 4.13  | 4.22       | 4.04       | 4.13       |
| TPR                | 3.65  | 3.76       | 3.57       | 3.60       |
| Difference between SR vs. TPR | 0.49  | 0.46       | 0.47       | 0.53       |
| ($t$; $p$)          | (20.46; 0.00) | (16.40; 0.00) | (16.83; 0.00) | (17.84; 0.00) |

As predicted by H3a, both SR and TPR showed nomological validity by positively correlating with both work condition and living conditions (Table 5). Furthermore, the differences in the satisfaction–engagement correlation between SR and TPR was moderated by whether the condition was good or poor. The difference was bigger when the environment was poor. For work environment, if the condition was poor, the difference between SR and TPR was 0.58, and 0.38 if the condition was good. Similarly, for living environment, if the condition was poor, the difference between SR and TPR was 0.59, and 0.38 if the condition was good (Table 5). H3b was supported.

Table 5. Correlations between environment satisfactions and overall work engagement.

|                   | SR    | TPR   | Difference between SR and TPR | ($t$; $p$) |
|-------------------|-------|-------|-------------------------------|------------|
| Work condition    |       |       |                               |            |
| Poor              | 3.89  | 3.31  | 0.58                          | (18.23; 0.00) |
| Good              | 4.39  | 4.01  | 0.38                          |            |
| Living condition  |       |       |                               |            |
| Poor              | 3.91  | 3.33  | 0.59                          | (19.48; 0.00) |
| Good              | 4.34  | 3.96  | 0.38                          |            |

Hypothesis 4a stated that EWE measured by self-rating from male respondents is higher than that from female respondents. Table 6 shows the differences in mean EWE ratings between males and females and its three dimensions for both SR and TPR methods. For SR, although males have indicated
slightly higher mean engagement ratings (4.15) than females (4.12), the difference was not significant ($p = 0.56$). H4a was not supported by the result. The finding is consistent with that of Wilson [23], where she reported male participants to have a higher but nonsignificant mean score than female participants. A similar pattern exists for the three dimensions: vigor (4.25 vs. 4.19, $p = 0.19$), dedication (4.06 vs. 4.03, $p = 0.65$), and absorption (4.13 vs. 4.14, $p = 0.91$). These results are different from that of Schaufeli and Bakker [25]. Based on their total data, they reported that "... men (N = 6469) score significantly higher than women (N = 5722), on all three aspects of engagement: means for men on vigor, dedication, and absorption are 4.28, 3.83, and 4.36, respectively, against 4.11, 3.77, and 4.26 for women." (p. 32).

Also shown in Table 6, for TPR, females indicated significantly higher mean EWE ratings and its three dimensions than males (overall engagement: 3.71 vs. 3.58, $p < 0.05$; vigor: 3.82 vs. 3.69, $p < 0.05$; dedication: 3.64 vs. 3.47, $p < 0.01$; absorption: 3.66 vs. 3.54, $p = 0.09$). H4b was supported by the result. Furthermore, as shown in Table 7, the difference in mean EWE between SR and TPR was significantly higher for males than for females (0.57 for males, and 0.41 for females; $F = 11.71; p < 0.00$). The same pattern applied to the two of the three dimensions of engagement: vigor (0.56 vs. 0.38, $p < 0.00$); dedication (0.58 vs. 0.39, $p < 0.00$); and, with one exception dimension: absorption (0.59 vs. 0.48, $p = 0.07$). H4c is confirmed that self-rating of EWE is higher than third-person rating, significantly more so for males than for females. In short, gender affects EWE if TPR used, but not relevant if SR is used.

H5a stated that the longer an employee works in a company, the lower EWE level will be. As shown in Table 8, between those who work for less than a year and those who work one to two years, the EWE level dipped slightly, but not significantly. However, contrary to the expectation of H5a, as one passed the first two years of engagement, participants expressed a significantly ($p < 0.01$) higher level of EWE as the length of job tenure increased. The pattern was true for both SR and TPR, and for the overall engagement level, as well as for the three dimensions, with one exception (vigor measured by TPR was not significant; $p = 0.06$). Wilson [23] also found a nonsignificant ($p = 0.12$), yet positive relationship, between the level of engagement and years of service. H5a was not supported. In addition, as shown in Table 9, the difference in EWE between SR and TPR significantly increased as the length of job tenure increased ($F = 3.02, p = 0.02$). For the three dimensions, the same pattern applied to dimension of engagement: vigor ($F = 3.27, p = 0.01$), while the difference between SR and TPR in the dimensions of dedication ($F = 2.07, p = 0.08$) and absorption ($F = 2.30, p = 0.06$) were a little higher than the recommended significant level ($p < 0.5$). Thus, H5b was supported.
Table 8. Years of tenure and engagement ratings.

| Length of Working at Company | Overall Engagement | Vigor | Dedication | Absorption |
|-----------------------------|-------------------|-------|------------|------------|
|                            | SR    | TPR  | SR    | TPR  | SR    | TPR  | SR    | TPR  |
| <1 Year                     | 3.98  | 3.60 | 4.09  | 3.78 | 3.84  | 3.51 | 4.01  | 3.47 |
| n = 125                     |       |      | n = 125|      | n = 125|      | n = 125|      |
| 1–2 Years                   | 3.95  | 3.49 | 4.11  | 3.67 | 3.82  | 3.36 | 3.91  | 3.43 |
| n = 118                     |       |      | n = 118|      | n = 118|      | n = 118|      |
| 3–5 Years                   | 4.05  | 3.61 | 4.11  | 3.70 | 4.00  | 3.51 | 4.05  | 3.63 |
| n = 155                     |       |      | n = 155|      | n = 155|      | n = 155|      |
| 6–9 Years                   | 4.17  | 3.63 | 4.26  | 3.68 | 4.11  | 3.60 | 4.18  | 3.59 |
| n = 94                      |       |      | n = 94 |      | n = 94 |      | n = 94|      |
| ≥10 Years                   | 4.44  | 3.83 | 4.49  | 3.91 | 4.37  | 3.78 | 4.46  | 3.80 |
| n = 148                     |       |      | n = 148|      | n = 148|      | n = 148|      |

\[(F, p)\] (13.26; 0.00) (3.50; 0.01) (10.35; 0.00) (2.23; 0.06) (9.27; 0.00) (4.05; 0.00) (11.36; 0.00) (3.47; 0.01)

Table 9. Years of tenure and differences in means between SR and TPR.

| Length of Working at Company | Overall Engagement (SR-TPR) | Vigor (SR-TPR) | Dedication (SR-TPR) | Absorption (SR-TPR) |
|-----------------------------|-------------------------------|---------------|---------------------|---------------------|
| <1 year                     | 0.39                          | 0.32          | 0.34                | 0.54                |
| 1–2 years                   | 0.45                          | 0.45          | 0.45                | 0.46                |
| 3–5 years                   | 0.44                          | 0.41          | 0.48                | 0.41                |
| 6–9 years                   | 0.56                          | 0.58          | 0.51                | 0.58                |
| ≥10 years                   | 0.61                          | 0.58          | 0.58                | 0.65                |
| \[(F, p)\]                  | (3.02; 0.02)                  | (3.27; 0.01)  | (2.07; 0.08)        | (2.30; 0.06)        |

Finally, Table 10 shows the relationship between the level of engagement and the position held by the respondents. There was a significant \((p < 0.00)\) and positive relationship between the level of engagement and the job position. When the frontline service staff, frontline operation staff, and maintenance personnel were collapsed together into one nonsupervisory workers category, the pattern became even clearer. Respondents with management responsibility (supervisors included) expressed a higher engagement level. Within the management groups, the higher the position, the higher the engagement level. For example, Table 11 shows that the mean SR engagement for nonsupervisory workers was 3.99; supervisors, 4.16; mid-level managers, 4.52; and upper-level managers, 4.64. There was a marked jump in engagement from supervisors to mid-level managers. The pattern was true for the three dimensions of engagement as well. It also applied to both SR and TPR. H6a was supported and the results indicate that managers and supervisors reported higher levels of engagement than the non-managers/supervisors. Table 11 shows the difference in the mean engagement score between self-rating and third-person rating for employees with leadership positions versus those without leadership position. It is clear that the differences between mean scores from SR and TPR became bigger as employees move from frontline positions to supervisors, then to mid-management position, and ultimately top-management position \((F = 3.25, p < 0.05)\). H6b was supported. Overall, both SR and TPR showed the same pattern of relationships between leadership position and engagement; that is, managers and supervisors reported higher levels of engagement than the non-managers/supervisors.
Table 10. Job position and engagement ratings.

| Job Position     | Overall Engagement | Vigor | Dedication | Absorption |
|------------------|--------------------|-------|------------|------------|
|                  | SR TPR             |       | SR TPR     | SR TPR     | SR TPR     | SR TPR     |
| FLSS             | 3.90 3.56          | 4.01 3.70 | 3.79 3.47 | 3.92 3.49  |
|                  | n = 216 n = 216  | n = 215 | n = 215    | n = 215    | n = 215    |
| FLOS             | 4.17 3.69          | 4.24 3.79 | 4.07 3.59 | 4.20 3.70  |
|                  | n = 77 n = 77     | n = 78 | n = 77     | n = 78     | n = 77     |
| MP               | 4.03 3.52          | 4.26 3.72 | 3.79 3.30 | 4.04 3.54  |
|                  | n = 67 n = 67     | n = 67 | n = 67     | n = 67     | n = 67     |
| Supervisors      | 4.16 3.56          | 4.21 3.65 | 4.10 3.49 | 4.14 3.54  |
|                  | n = 147 n = 147   | n = 147 | n = 147    | n = 147    | n = 147    |
| M-L Managers     | 4.52 3.94          | 4.56 4.00 | 4.52 3.91 | 4.47 3.90  |
|                  | n = 75 n = 75     | n = 75 | n = 75     | n = 75     | n = 75     |
| U-L Managers     | 4.64 4.07          | 4.62 4.07 | 4.71 4.14 | 4.58 3.97  |
|                  | n = 25 n = 25     | n = 25 | n = 25     | n = 25     | n = 25     |
| (F; p)           | (12.34; 0.00)     | (10.00; 0.00) | (8.01; 0.00) | (5.74; 0.00) | (2.93; 0.01) |

Notes: FLSS = frontline service staff; FLOS = frontline operation staffs (i.e., cook, dishwasher); M-L Manager = mid-level managers; U-L Managers = upper-level managers.

Table 11. Collapsed job positions and differences in engagement ratings between SR and TPR.

| Job Positions   | Overall Engagement SR–TPR | Overall Engagement (SR–TPR) | Vigor (SR–TPR) | Dedication (SR–TPR) | Absorption (SR–TPR) |
|-----------------|---------------------------|-----------------------------|---------------|---------------------|---------------------|
| Frontline       | 3.99 3.58                 | 0.41                        | 0.39          | 0.39                | 0.46                |
| Supervisors     | 4.16 3.56                 | 0.60                        | 0.57          | 0.62                | 0.60                |
| M-L Managers    | 4.52 3.94                 | 0.58                        | 0.56          | 0.62                | 0.57                |
| U-L Managers    | 4.64 4.07                 | 0.57                        | 0.55          | 0.57                | 0.61                |
| (F; p)          | (19.84; 0.00)             | (7.57; 0.00)                | (3.25; 0.00)  | (2.32; 0.06)        | (3.99; 0.00)        |

Thus, ten out of twelve hypotheses have been supported, and two hypotheses were not supported (one was not significant and one showed the reverse result), which are summarized in Table 12.

Table 12. The results of hypotheses testing.

| Hypotheses          | Significance | Decision |
|---------------------|--------------|----------|
| Hypothesis 1 (H1)   | sig.         | Supported |
| Hypothesis 2 (H2a)  | sig.         | Supported |
| Hypothesis 2 (H2b)  | sig.         | Supported |
| Hypothesis 3 (H3a)  | sig.         | Supported |
| Hypothesis 3 (H3b)  | sig.         | Supported |
| Hypothesis 4 (H4a)  | n.s.         | Not supported |
| Hypothesis 4 (H4b)  | sig.         | Supported |
| Hypothesis 4 (H4c)  | sig.         | Supported |
| Hypothesis 5 (H5a)  | sig. Reverse result | |
| Hypothesis 5 (H5b)  | sig.         | Supported |
| Hypothesis 6 (H6a)  | sig.         | Supported |
| Hypothesis 6 (H6b)  | sig.         | Supported |

Notes: sig. = statistically significant; n.s. = statistically insignificant.
6. Discussion

Enhancing employee engagement is an important task for the implementation of SHRM [80]. Employees with higher levels of engagement have higher levels of satisfaction, commitment, performance, and low level of turnover intentions [81], which leads to sustainable workforce productivity. Thus, it is imperative to accurately measure employees’ level of engagement. SR has been predominantly used to assess the level of work engagement for employees. However, the SR method has also been known for suffering from ego-protective and ego-enhancing biases, potentially resulting false inflationary engagement level. It is only human nature to avoid providing self-incriminating evidences. To alleviate the inflationary bias of SR method, the present study investigated an alternative method to assess the EWE construct—the third-person rating (TPR) method. This study compared SR and TPR in a hospitality industry setting.

In the study, we were able to confirm that the three-factor structure developed for SR method fit pretty well for TPR method. The results also showed that TPR is highly correlated with SR, offering evidence of convergent validity. Furthermore, TPR shows a nomological validity by positively correlating with one of the antecedent variables of work engagement—satisfaction with work conditions and living condition. Most importantly, TPR, as an alternative measure, and SR of EWE have very similar relationships with their common correlates (satisfaction with environmental conditions; rater’s gender, length of work tenure, and level of leadership position, etc.). The patterns of relationships that SR and TPR EWE have with their common correlates are almost identical, and SR of EWE has a higher mean level of ratings. The only exception is the relationship with gender, where SR shows no relationship while TPR shows a relationship. A summary of specific findings and implications are as follows:

First, the finding that SR has a positive bias is consistent with the majority of research findings on this issue [46]. This positive bias is gender neutral, applying to both genders. Hence, the findings that the mean of SR of EWE was significantly higher than TPR of EWE signaled the existence of the “better-than-average” positive bias associated with the SR. The research also found that males not only “exaggerate” their own scores, they are also more “critical” of others’ performances. This is consistent with findings of Wilson [23] and Shukla et al. [82] and Khodakarami and Dirani [83]. Therefore, if the TPR method is to be used to collect engagement data, a “negative bias adjustment” may need to be applied to male participants. In the current study, the negative bias for males was about 0.13 out of a five-point Likert scale. Furthermore, the differences between SR and TPR are bigger when the work and living conditions are poor.

Second, contrary to H6a, this study found that after the first two years at work, the length of job tenure correlated positively with the level of engagement. Wilson [23] found a similar positive, yet, nonsignificant correlation between engagement and years of service. Perhaps the results reflected a self-selection effect. Those employees who were not engaged or who could not take the work challenge in the first two years were either fired or left voluntarily. Those who stayed after two years tended to be those who did enjoy the work and performed well. Once an employee is over the hump, he/she will more likely stay engaged and productive. As such, this is consistent with the findings that there is a strong negative relation between engagement and intention to leave [81].

Third, consistent with the findings of Wilson [23], this study also found managers and supervisors reported higher levels of engagement than did the non-managers/supervisors, whether it is measured by SR or by TPR. It is crucial that the management team avoid “overgeneralizing” their own positive engagement to all levels of employees, particularly the frontline workers, and fail to develop an empathy for the frontline workers. Management should realize that the work conditions and compensation package are far superior for the management positions than for the frontline workers.

Fourth, the findings that both SR and TPR exhibit same three-factor structure, convergent validity and nomological validity, and that both methods of EWE have a very similar relationships with their common correlates, support the view that both methods are viable ways to measure EWE. This is consistent with the call for using multiple methods to establish validity (multitrait–multimethod matrix
approach) advocated by researchers in hospitality sector [48]. Hence, hotel owners and managers should adopt multiple methods to accurately obtain employee engagement data rather than a single SR approach. However, they need to realize that TPR method does result in a substantially lower EWE score than that assessed by SR. Given the social desirability nature of the EWE construct, TPR result is more likely to be close to the true EWE level. Furthermore, committing a false positive (falsely indicating the presence of a condition, such as low engagement level, when it is not) is less costly than committing a false negative.

Furthermore, the mean differences between SR and TPR of EWE were found to be affected by the specific dimension under study, rater’s leadership position, work tenure, and work environment conditions. Specifically, since some dimensions of engagement are often subtle and difficult to judge from an outside viewpoint, how much the TPR method is subjected to the influence of negative bias may depend on the dimension of engagement one is measuring. Vigor reflects energy level and is easier to observe objectively by others. On the other hand, absorption indicates an inner mental state and is highly unobservable to others. As a result, the difference between SR and TPR is smallest for the vigor dimension and largest for the absorption dimension. The present study also found that when raters are males, hold lower or no leadership positions, and work conditions are perceived to be poor, the mean differences between SR and TPR of EWE are larger. If management intends to gauge the true level of employee engagement, and only one method is to be employed, then, TPR method should be a preferred method over the SR method to prevent being deceived into believing the inflated engagement scores. This is particularly true if the majority of the employees are males, the working conditions of the firm are poor, individual dimension scores as opposed to overall EWE score are to be analyzed, and frontline workers are the subjects of measurement.

Finally, for academics, the results of the study also suggest that some inconsistent results from different studies may exist at measurement level, rather than at the underlying constructs level. They are caused by the different measurement methods used in various studies, particularly if the constructs involved have social desirability characteristics.

7. Conclusions

The purpose of the paper was to present the survey findings of two alternative methods (self-rating, SR, and third-person rating, TPR) of measuring employee work engagement (EWE). An accurate measure of employee work engagement serves as a leading indicator of turnover intention and an early diagnostic tool for sustainable human resource management. Despite its popularity as a work engagement measure, the SR method has many limitations. This research attempted to demonstrate that TPR is a viable and better alternative measure of EWE. The results indicated that TPR does possess desirable measurement characteristics, such as convergent validity, nomological validity, and structure invariant. TPR also provides a more conservative, and perhaps more accurate as well, measure of EWE. Our findings confirm that SR and TPR of EWE generally result in very similar patterns of findings, however, with the means of SR significantly higher than that obtained from TPR. The one exception is gender. Gender does not affect EWE score under SR, but exerts a significant impact under TPR. The mean of TPR engagement score for males does drop significantly greater than for females, indicating a self-“exaggerating” effect by males. Given that management is interested in not only the relationships but sometimes more importantly the level of employee work engagement, TPR does provide a more conservative, and perhaps more accurate as well, measure of EWE. A low EWE scores among employees provides a clear warning sign to the management that remedy actions are in order.

The importance of this study lies in the fact that the measurement of employee work engagement is one of the key steps to improve employee engagement and enhance sustainable HRM. For academics, the results in the study that SR and TPR of EWE generally result in similar pattern of findings offers strong encouragement to build future research on the TPR method. For practitioners, the importance of this study lies in the fact that as a conservative measure of EWE, TPR can play an important role
in detecting early sign of employee troubles sooner and lead management to take timely actions, making human resource management more sustainable.

While this study made some valuable findings for the hospitality sector, future research is still needed to reveal more detail information on hotel employee engagement. For example, more attention should be paid on the young generation [84] or the short-term employees (less than two years). As most of the employees who directly provide services to customers are frontline employees [85], future research can also focus on determining whether TPR is the most appropriate method to measure the EWE of the frontline employees. Finally, although a representative sample meeting the quantitative criteria was used in this paper, future studies might obtain data from enterprises in various regions/countries to verify the results of this study. There is also a need to further explore the characteristics of TPR as a viable measure method by investigating the predictive validity of TPR on some outcome variables, such as retention rate, job satisfaction rate, etc. It will advance our knowledge about the boundary of TPR by comparing SR with TPR when the information sought is not self-incriminating or is not easily observable, such as motives for work engagement.

**Author Contributions:** Conceptualization, F.Z.X., H.Y., and B.T.W.; data curation, B.T.W.; formal analysis, B.T.W. and Y.Z.; funding acquisition, H.Y. and F.Z.X.; investigation, F.Z.X.; methodology, F.Z.X., B.T.W., and Y.Z.; project administration, H.Y.; supervision, H.Y.; validation, B.T.W. and Y.Z.; writing—original draft, F.Z.X., B.T.W., H.Y., and Y.Z.; writing—review and editing, F.Z.X., Y.Z., H.Y., and B.T.W. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by Key Projects of the National Social Science Fund of China (18AJY011), Shandong Natural Science Foundation of China (ZR2019MG014), and Teaching Cases of Professional Master Degree of Shandong Province in 2019 (SDYAL19010).

**Conflicts of Interest:** The authors declare no conflict of interest.

**References**

1. Ting, C.-T.; Hsieh, C.-M.; Chang, H.-P.; Chen, H.-S. Environmental Consciousness and Green Customer Behavior: The Moderating Roles of Incentive Mechanisms. *Sustainability* **2019**, *11*, 819. [CrossRef]

2. Saura, J.R.; Reyes-Menendez, A.; Alonso, C. Álvarez Do Online Comments Affect Environmental Management? Identifying Factors Related to Environmental Management and Sustainability of Hotels. *Sustainability* **2018**, *10*, 3016. [CrossRef]

3. United Nations. Department of Economic and Social Affairs: Sustainable Development. Available online: [https://sdgs.un.org/goals](https://sdgs.un.org/goals) (accessed on 1 August 2020).

4. Ehnert, I.; Parsa, S.; Roper, I.; Wagner, M.; Müller-Camen, M. Reporting on sustainability and HRM: A comparative study of sustainability reporting practices by the world’s largest companies. *Int. J. Hum. Resour. Manag.* **2015**, *27*, 88–108. [CrossRef]

5. Cain, L.; Busser, J.; Kang, H.J. Executive chefs’ calling: Effect on engagement, work-life balance and life satisfaction. *Int. J. Contemp. Hosp. Manag.* **2018**, *30*, 2287–2307. [CrossRef]

6. Kaya, B.; Karatepe, O.M. Does servant leadership better explain work engagement, career satisfaction and adaptive performance than authentic leadership? *Int. J. Contemp. Hosp. Manag.* **2020**, *32*, 2075–2095. [CrossRef]

7. Kahn, W. The essence of employee engagement: Lessons from the field. In *Handbook of Employee Engagement: Perspectives, Issues, Research, and Practice*; Albrecht, S.L., Ed.; Edward Elgar: Northampton, MA, USA, 2010; pp. 20–30.

8. Maslach, C.; Schaufeli, W.B.; Leiter, M.P. Job Burnout. *Annu. Rev. Psychol.* **2001**, *52*, 397–422. [CrossRef]

9. Saks, A.M. Antecedents and consequences of employee engagement. *J. Manag. Psychol.* **2006**, *21*, 600–619. [CrossRef]

10. Ariza-Montes, A.; Arjona-Fuentes, J.M.; Han, H.; Law, R. Work environment and well-being of different occupational groups in hospitality: Job Demand–Control–Support model. *Int. J. Hosp. Manag.* **2018**, *73*, 1–11. [CrossRef]

11. Kim, C.Y.; Lee, J.H.; Shin, S.Y. Why Are Your Employees Leaving the Organization? The Interaction Effect of Role Overload, Perceived Organizational Support, and Equity Sensitivity. *Sustainability* **2019**, *11*, 657. [CrossRef]

12. Moyle, B.D.; Moyle, B.; Ruhanen, L.M.; Bec, A.; Weiler, B. Business Sustainability: How Does Tourism Compare? *Sustainability* **2018**, *10*, 968. [CrossRef]
13. Szromek, A.R.; Kruczek, Z.; Walas, B. The Attitude of Tourist Destination Residents towards the Effects of Overtourism—Kraków Case Study. *Sustainability* 2019, 12, 228. [CrossRef]
14. Trišić, I.; Stetić, S.; Privitera, D.; Nedelcu, A. Wine Routes in Vojvodina Province, Northern Serbia: A Tool for Sustainable Tourism Development. *Sustainability* 2019, 12, 82. [CrossRef]
15. Shen, L.; Qian, J.; Chen, S.C. Effective Communication Strategies of Sustainable Hospitality: A Qualitative Exploration. *Sustainability* 2020, 12, 6920. [CrossRef]
16. Martin, M.; Angel, R.; Folgado-Fernández, J.A.; Palos-Sanchez, P.; Castejón-Jiménez, P. The Impact of the Environmental Quality of Online Feedback and Satisfaction When Exploring the Critical Factors for Luxury Hotels. *Sustainability* 2019, 12, 299. [CrossRef]
17. Trafiałek, J.; Czarniecka-Skubina, E.; Kulaitienė, J.; Vaitkevičienė, N. Restaurant’s Multidimensional Evaluation Concerning Food Quality, Service, and Sustainable Practices: A Cross-National Case Study of Poland and Lithuania. *Sustainability* 2019, 12, 234. [CrossRef]
18. Wikhamn, W. Innovation, sustainable HRM and customer satisfaction. *Int. J. Hosp. Manag.* 2019, 76, 102–110. [CrossRef]
19. 8 Employee Engagement Statistics You Need to Know in 2020. 9 January 2020. Available online: https://www.gallup.com/workplace/257552/state-global-workplace-2017.aspx (accessed on 1 July 2020).
20. State of the Global Workplace. 14 May 2019. Available online: https://www.gallup.com/workplace/257552/state-global-workplace-2017.aspx (accessed on 1 July 2020).
21. De Cuyper, N.; Schreurs, B.; Elst, T.V.; Baillien, E.; De Witte, H. Exemplification and Perceived Job Insecurity. *J. Pers. Psychol.* 2014, 13, 1–10. [CrossRef]
22. Stewart, S.M.; Bing, M.N.; Davison, H.K.; Woehr, D.J.; McIntyre, M.D. In the eyes of the beholder: A non-self-report measure of workplace deviance. *J. Appl. Psychol.* 2009, 94, 207–215. [CrossRef]
23. Wilson, K. A Survey of Employment Engagement. Ph.D. Thesis, The Faculty of the Graduate School, University of Missouri-Columbia, Columbia, MO, USA, 2009.
24. Podsakoff, P.M.; MacKenzie, S.B.; Podsakoff, N.P. Sources of Method Bias in Social Science Research and Recommendations on How to Control It. *Annu. Rev. Psychol.* 2012, 63, 539–569. [CrossRef]
25. Schaufeli, W.B.; Bakker, A.B. UWES Utrecht Work Engagement Scale: Test manual. Utrecht University. 2003. Available online: www.Schaufeli.com (accessed on 20 October 2019).
26. Ehnert, I.; Harry, W. Recent Developments and Future Prospects on Sustainable Human Resource Management: Introduction to the Special Issue. *Manag. Rev.* 2012, 23, 221–238. [CrossRef]
27. Bondarouk, T.; Brewster, C.J. Conceptualising the future of HRM and technology research. *Int. J. Hum. Resour. Manag.* 2016, 27, 2652–2671. [CrossRef]
28. Macke, J.; Genari, D. Systematic literature review on sustainable human resource management. *J. Clean. Prod.* 2019, 208, 806–815. [CrossRef]
29. Macey, W.; Schneider, B. The meaning of employment engagement. *Ind. Organ. Psychol.* 2008, 1, 3–30. [CrossRef]
30. Kahn, W.A. Psychological conditions of personal engagement and disengagement at work. *Acad. Manag. J.* 1990, 33, 321–349.
31. Rich, B.L.; Lepine, J.A.; Crawford, E.R. Job Engagement: Antecedents and Effects on Job Performance. *Acad. Manag. J.* 2010, 53, 617–635. [CrossRef]
32. Byrne, Z.S.; Peters, J.M.; Weston, J.W. The struggle with employee engagement: Measures and construct clarification using five samples. *J. Appl. Psychol.* 2016, 101, 1201–1227. [CrossRef] [PubMed]
33. Shuck, B.; Adelson, J.L.; Reio, T.G. The Employee Engagement Scale: Initial Evidence for Construct Validity and Implications for Theory and Practice. *Hum. Resour. Manag.* 2016, 56, 953–977. [CrossRef]
34. González-Romá, V.; Schaufeli, W.B.; Bakker, A.B.; Lloret, S. Burnout and work engagement: independent factors or opposite poles? *J. Vocat. Behav.* 2006, 68, 165–174. [CrossRef]
35. Joseph, D.L.; Newman, D.A.; Hulin, C.L. Job Attitudes and Employee Engagement: A Meta-Analysis of Construct Redundancy. In *Academy of Manag. Proceedings*; Academy of Management: New York, NY, USA, 2010; Volume 2010, pp. 1–6.
36. Brayfield, A.H.; Rothe, H.F. An index of job satisfaction. *J. Appl. Psychol.* 1951, 35, 307–311. [CrossRef]
37. Christian, M.; Garza, A.S.; Slaughter, J.E. Work Engagement: A Quantitative Review and Test of its Relations with Task and Contextual Performance. *Pers. Psychol.* 2011, 64, 89–136. [CrossRef]
38. Baran, M.; Sypniewska, B. The Impact of Management Methods on Employee Engagement. *Sustainability* 2020, 12, 426. [CrossRef]
39. Jung, H.S.; Yoon, H.H. How Does Workplace Romance Influence Employee Performance in the Hospitality Industry? *Sustainability* 2020, 12, 5478. [CrossRef]
40. Grobelna, A. Effects of individual and job characteristics on hotel contact employees’ work engagement and their performance outcomes. *Int. J. Contemp. Hosp. Manag.* 2019, 31, 349–369. [CrossRef]
41. Miao, S.; Rhee, J.; Jun, I. How Much Does Extrinsic Motivation or Intrinsic Motivation Affect Job Engagement or Turnover Intention? A Comparison Study in China. *Sustainability* 2020, 12, 3630. [CrossRef]
42. Gigol, T. Influence of Authentic Leadership on Unethical Pro-Organizational Behavior: The Intermediate Role of Work Engagement. *Sustainability* 2020, 12, 1182. [CrossRef]
43. Podsakoff, M.; MacKenzie, S.; Lee, J.Y. Common method biases in behavioral research: A review of the literature and recommended remedies. *J. Appl. Psychol.* 2003, 88, 879–903. [CrossRef]
44. Ng, T.W.; Feldman, D.C. A comparison of self-ratings and non-self-report measures of employee creativity. *Hum. Relat.* 2012, 65, 1021–1047. [CrossRef]
45. Fisher, R.J. Social Desirability Bias and the Validity of Indirect Questioning. *J. Consum. Res.* 1993, 20, 303–315. [CrossRef]
46. Guilford, J.P. *Psychometric Methods*, 2nd ed.; McGraw-Hill: New York, NY, USA, 1954.
47. Berry, C.M.; Carpenter, N.C.; Barratt, C.L. Do other-reports of counterproductive work behavior provide an incremental contribution over self-reports? A meta-analytic comparison. *J. Appl. Psychol.* 2012, 97, 613–636. [CrossRef]
48. Alicke, M.D.; Klotz, M.L.; Breitenbecher, D.L.; Yurak, T.J.; Vredenburg, D.S. Personal contact, individuation, and the better-than-average effect. *J. Personal. Soc. Psychol.* 1993, 68, 804–825. [CrossRef]
49. Rosenthal, S.; Detenber, B.H.; Rojas, H. Effects of functional green advertising on self and others. *Psychol. Mark.* 2019, 34, 141–150. [CrossRef]
66. Shore, L.M.; Thornton, G.C. Effects of gender on self and supervisory ratings. *Acad. Manag. J.* 1986, 29, 115–129.
67. Magee, W. Effects of Gender and Age on Pride in Work, and Job Satisfaction. *J. Happiness Stud.* 2014, 16, 1091–1115. [CrossRef]
68. Sturm, R.E.; Taylor, S.N.; Atwater, L.E.; Braddy, P.W. Leader self-awareness: An examination and implications of women’s under-prediction. *J. Organ. Behav.* 2014, 35, 657–677. [CrossRef]
69. Keller, E. Reflections on Gender and Science; Yale University Press: New Haven, NH, USA, 1985.
70. Maccoby, E.E. Gender and relationships. *Am. Psychol.* 1990, 45, 513–520. [CrossRef]
71. Coffman, C.; Gonzalez-Molina, G. Follow this Path: How the World’s Greatest Organizations Drive Growth by Unleashing Human Potential; Warner Books: New York, NY, USA, 2002.
72. Attridge, M. Measuring and Managing Employee Work Engagement: A Review of the Research and Business Literature. *J. Work. Behav. Health* 2009, 24, 383–398. [CrossRef]
73. Kerfoot, K.M. Staff engagement: It starts with the leader. *Medsurg Nurs. Off. J. Acad. Med.-Surg. Nurses* 2008, 17, 64–65.
74. Hambleton, R.K. Translating achievement tests for use in cross-national studies. *Eur. J. Psychol. Assess.* 1993, 9, 57–65.
75. Attridge, M. Measuring and Managing Employee Work Engagement: A Review of the Research and Business Literature. *J. Appl. Psychol.* 2005, 90, 1217–1227. [CrossRef]
76. Lee, Y.; Chang, H. Relations between team work and innovation in organizations and the job satisfaction of employees: A factor analytic study. *Int. J. Manag.* 2008, 25, 732–739.
77. Hair, J.F.; Black, W.C.; Babin, B.J.; Anderson, R.E. *Multivariate Data Analysis*, 7th ed.; Pearson Education: Upper Saddle River, NJ, USA, 2010.
78. Hair, J.F.; Black, W.C.; Babin, B.J.; Anderson, R.E. *Multivariate Data Analysis*, 7th ed.; Pearson Education: Upper Saddle River, NJ, USA, 2010.
79. Shukla, S.; Adhikari, B.; Singh, V. Employee engagement-role of demographic variables and personality factors. *Amity Glob. HRM Rev.* 2015, 5, 65–73.
80. Khodakarami, N.; Dirani, K. Drivers of employee engagement: Differences by work area and gender. *Ind. Commer. Train.* 2020, 52, 81–91. [CrossRef]
81. Cismaru, L.; Iunius, R. Bridging the Generational Gap in the Hospitality Industry: Reverse Mentoring—An Innovative Talent Management Practice for Present and Future Generations of Employees. *Sustainability* 2019, 12, 263. [CrossRef]
82. Xu, F.Z.; Wang, Y. Enhancing Employee Innovation through Customer Engagement: The Role of Customer Interactivity, Employee Affect, and Motivations. *J. Hosp. Tour. Res.* 2019, 44, 351–376. [CrossRef]