Social Indicators Research (2019) 146:61–76
https://doi.org/10.1007/s11205-019-02071-6

How Do Students Evaluate Instructors' Performance? Implication of Teaching Abilities, Physical Attractiveness and Psychological Factors

Sharon Tan[1] · Evan Lau[2] · Hiram Ting[3,4] · Jun-Hwa Cheah[5] · Biagio Simonetti[6] · Hiok Lip Tan[2]

Accepted: 16 January 2019 / Published online: 15 February 2019

Abstract
One instrument regularly seen as a basic resource in assessing pedagogical knowledge and vivid learning in different circumstances is through the method of conducting student assessment appraisal of their instructors. Nevertheless, deciding the nature of instructional abilities requires as rationale and unbiased judgments. The concern is that there are no formal techniques or formulas that would prompt accurate responses from the students. In spite of the contention surrounding students' rating on instructors, this study aims to investigate how university students in Malaysia would evaluate instructors based on non-instructional factors, such as physical attractiveness and psychological factors, which in turn may affect students' perceptions towards instructors' performance. PLS-SEM was appropriated to perform the path modeling analysis. Practical implication is discussed.

Keywords Evaluation of instructors' performance · Physical attractiveness · Teaching abilities

JEL Clasification I20 · I21 · J24

✉ Evan Lau
lphevan@unimas.my; lphevan@gmail.com

[1] Faculty of Accountancy, Finance and Business, Tunku Abdul Rahman University College, 89500 Penampang, Sabah, Malaysia

[2] Faculty of Economics and Business, Universiti Malaysia Sarawak, 94300 Kota Samarahan, Sarawak, Malaysia

[3] Faculty of Hospitality and Tourism Management, UCSI University, 93050 Kuching, Sarawak, Malaysia

[4] Sarawak Research Society, 94200 Kuching, Sarawak, Malaysia

[5] Department of Management and Marketing, Faculty of Economics and Management, Universiti Putra Malaysia (UPM), 43400 Serdang, Kuala Lumpur, Selangor, Malaysia

[6] Department of Law, Economics, Management and Quantitative Methods, University of Sannio, Via delle Puglie, 8282100 Benevento, Italy

1 Introduction

One of the many challenges faced by universities today is to improve instructors' teaching and other relevant educational abilities in order to enhance students' learning process in a conducive manner. Ironically, the indelibility of an instructor's teaching ability has unprecedented implications especially in the quest to emphasize on educational objectives, stimulate students' interest, and enhance cognitive development in students' educational journey. The effectiveness of university teaching varies across a wide range of conditions due to diversity in the modes of instructions which include lecturing, practical lab sessions and mentoring. However, numerous studies on measuring instructors' teaching abilities have also highlighted potential biases which affect students' interest in learning and their subsequent verdicts of instructors' performance (Marsh 2007; Moss-Racusin et al. 2012; Ponzo and Scoppa 2013).

Like the concept of customers in business, students' experience and satisfaction in universities is pivotal to improving the quality of instructors' teaching abilities and the utilization of teaching methods which foster effective and long term involvement with learning tasks. A strong association also exists with ability to rely on the support of significant others, self-efficacy and social abilities, and sense of belonging to different significant social groups (McGrath et al. 2009). The educational expectations on one hand, and the goal commitment on the another hand, change during the time and can depend by social and structural integration into the systems of the institution (Ingusci et al. 2016).

In practice, universities have adopted this concept through the means of conducting students' evaluation assessment to seek out their inclination to approach learning and conditions that induce them to develop intellectual growth. Although it has been largely assumed that students would evaluate instructors' performance based on their teaching qualities, it is reported that non-instructional factors, such as physical attractiveness, may affect students' perceptions towards instructors' teaching abilities or credibility (Kogan et al. 2010). It is well documented that physical attractiveness is a powerful social tool and people who are more attractive receive preferential treatments as compared to others who are less (Hamermesh and Parker 2005; Olson and Marshuetz 2005; Gurung and Vespia 2007). As such, its potential effect on students' evaluation of instructors cannot be overlooked. Organizational decisions may be split into the broad categories of content and process. The effectiveness of strategic decision-making processes depends on multiple criteria. One of the key ingredients of effective decision-making is to be evidence-based (Kurtulmuş et al. 2016).

Since students' evaluation and feedback has become the key component to attaining quality in university, it is increasingly used to evaluate instructors' competencies and thus serves as a mechanism to determine the best practices in instructional management. Information about students' satisfaction inferred from the survey may be a good starting-point to begin a discussion between teachers and students about the concept of good teaching': students' evaluations could be analysed together, in order to understand the position of each of them, by sharing and comparing different points of view (Bassi et al. 2017). This could activate mechanisms of real involvement of the principal exponents of teaching and learning, by means of which they could experience new kinds of participation in university life and contribute to its changes.

All reforming efforts and the prevalent use of students' evaluation to improve instructions have led to the voluminous literature on the various aspects of providing high-quality learning experiences to motivate students to expend the effort to learn (Benton and Cashin

2014; Chan et al. 2012; Marsh 2007). Whilst the criterion to assess the instructors remains elusive, affective traits, such as degree of confidence and level of attention, cannot be easily determined. Albeit the growing body of knowledge in the principles and maxims of teaching, restructuring instructional objectives to maintain students' interest can be challenging, thus requiring immediate attentions (Spooren et al. 2013).

In light of the aforementioned, it is apparent that there is a need to look into the effect of instructors' teaching abilities and physical attractiveness on students' evaluation. Although the profound effect of non-instructional factors on human behaviour is irrefutable, there is an extreme lack of literature that articulates the impact of instructors' physical attractiveness on students' evaluation of instructors' performance in Malaysian universities. Moreover, the investigation pertaining to the effect of psychological factors, such as motivation, degree of confidence and level of attention on evaluation of instructors' performance in academia, which could potentially elucidate how students evaluate, are also found wanting. Hence, the present study is aimed to firstly determine the effect of teaching abilities and physical attractiveness of the instructors on students' motivation, degree of confidence and level of attention. Secondly, it is to ascertain the effect of the three psychological factors on students' evaluation of the instructors' performance. Thirdly, the three psychological factors would be tested as mediators to assess their effect on the relationship between the antecedents and performance evaluation.

2 Literature Review

2.1 Teaching Abilities and Physical Attractiveness

In past couple of decades, universities have paid more attention to improve instructors' teaching abilities and other relevant educational qualities. According to De Paola and Scoppa (2015), committee members are appointed by university to mainly evaluate lecturers' skills which include teaching abilities so as to choose and promote the right candidates to higher positions. In order to ensure continual development in teaching methods and materials, some universities establish centres or institutes to specifically deal with these objectives (Smimou and Dahl 2012). Understandably, when these centres and institutes evaluate the instructors, such as their strengths and weaknesses, their teaching abilities are always put to task.

Physical attractiveness, in turn, creates a physiological effect, builds self-confidence and enhances self-motivation and productivity. Self-confidence is a key factor to improving motivation to accept life's challenges, perform and produce better results, and effort in action to achieve goals (Cipriani and Zago 2011). In addition, people with physical attractiveness is associated with their self-ability, where they look smarter and gain more attention from others and are specially treated by others. They appear to be more professional, trained and skillful (Hatfield and Rapson 2000). This is supported by Mobius and Rosenblat (2006) who claim that preferential treatment enhances the confidence and social plus communication abilities. Past studies have also asserted that physical attractiveness enhances social communication ability and interpersonal relationship (as cited in Eagly et al. 1991). It is found that good physical impression on people plays positive role. Specifically, most young people have positive impression towards the good looking people (Dion et al. 1972). They think good looking people own various good characteristics, like polite, fun, interesting, more sociable, than average looking people.

As such, this study investigates whether teaching abilities and physical attractiveness of the instructors affect students' psychological orientations, which result in their evaluation of instructors' performance at universities. In fact, the relationship between physical attractiveness and performance of instructors, and subsequently students' evaluation of instructors is not something entirely new (Hamermesh and Parker 2005), but little is done to articulate this matter. The existence of physical attractiveness stereotype among students' perception towards instructors is evident as students tend to evaluate higher performance and give higher rate to their instructors who look smart (MacNell et al. 2015). In order to enhance the understanding of students' evaluation, psychological factors are included in the study and they are discussed in the following paragraphs.

2.2 Motivation, Degree of Confidence and Level of Attention

Motivation is a driving effort to take action and to transform imagination into reality (Ryan and Deci 2000). Physical attractiveness creates a physiological effect in self-confidence, and enhances self-motivation and productivity. One of the most important skills for a student to develop during their educative journey is the process of learning. Recognizing that motivation is a key factor in this process, instructors apply teaching practices that help students to adapt to novel situations (Fulmer and Turner 2014). This philosophy is also supported by previous literature addressing the issues in students' motivation which underscores the magnitude of the learning environment (Byrne and Flood 2005). As students are faced with various challenges, for example a misalignment of students' expectations in an education setting, lack of motivation and unpreparedness towards the approach of learning will affect the manner they evaluate things (Smimou and Dahl 2012).

Students' competency level in their learning activity increases when their instructors' teaching abilities are augmented by prior experiences and vast knowledge (Klassen and Chiu 2010; Putman 2012). Therefore, it is imperative that their instructors are evaluated as this acts as a stimulus to improve classroom performances. Typically, a person's confidence level is deemed to be an ability or skill that contributes to the establishment and reinforcement of that person's motivation, thereby enabling it to be used in multitudes of daily tasks (Bénabou and Tirole 2002). Consequently, this necessitates a ramp-up of students' confidence toward their instructors as this would motivate the students in their learning endeavour. Additionally, instructors who acquire high levels of self-efficacy in performing their tasks are believed to be able to benefit students' learning progression (Bakar et al. 2013). Students' performance in school would be influenced positively by better efficacy of their instructors as it was indicated that better efficacy relate positively to performance outcomes (Tschannen-Moran and Hoy 2001).

One of the widespread justifications of using student's feedback for staff appraisal is largely due to the need to address deficiencies in pedagogic skills (Spooren et al. 2013). It is believed that students' poor evaluation of instructors' teaching competencies is attributed to their concern with the instructors' relevant abilities to conduct classroom activities. Therefore given this scenario, it is both physically and emotionally draining for the instructors to maintain interest and level of attention from the students especially with the challenge to balance between students' focus and transmitting the fundamental knowledge of a discipline (Copland et al. 2009). Ironically, there are other elements which could influence teaching performances other than interest and engagement in a student's learning process (Chan et al. 2012). As such, it underscores the relevance of

students' level of attention with instructors' teaching abilities and personalities, which would in turn enhance instructors' performance and favourable evaluation accordingly (Eison 1990).

2.3 Evaluation of Instructors' Performance

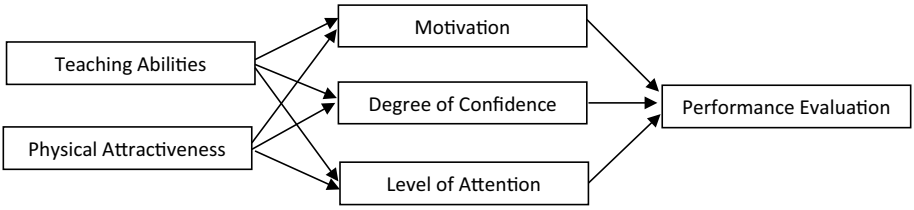
With the increase demand of higher institutions to achieve global competitiveness, conscious efforts to regulate, coordinate and assure quality remain as key determinants in achieving acceptable standards of education, scholarship, and infrastructure (Coates 2005). Whilst there are various factors affecting the quality in these institutions, the talent and expertise of teaching staff remains instrumental in reflecting students' competence in disciplinary content (Johnson 2012). In order to create a conducive learning environment, students' evaluation on instructors' performance are conducted on a yearly basis to allow provision for benchmarking purposes. Although the literature demonstrated that good teaching encompasses the contents of what is taught and other situational factors, nevertheless attention should not focus on a lecturers pedagogical skills, but also include the learning environment that is benevolent towards students personal needs (Mousavi et al. 2015).

In more recent paper by Simonetti et al. (2017) results showed that the influence of the influence of Organization and Infrastructures (OS) on the Student Satisfaction (SS) was indirect through the Didactics. In this sense didactics bridge the link between OS and SS for the university students in University Federico II of Naples, Italy.

In using these evaluations, the main outcome from such feedback is to properly plan improvements in order to overcome the challenges surrounding the quality of a student learning experiences. There is still much debate on how these evidences collected from students evaluation can be used to find mechanistic ways to improve an instructors teaching ability as the willingness and capacity to act upon the information provided is still questionable (Barrie and Ginns 2007). Given this backdrop, it is noteworthy to point out that much research that relates to the characteristics of the instructor has been done and the variables tested includes instructor rank and experience, the reputation and research skill of the instructor, along with more base concerns of gender, minority status and physical appearance (Worthington 2002). This current study will attempt to test these purported factors in a Malaysian university with the intention to extend empirical work in this area by investigating a student's characteristics and variables that are unrelated to effective teaching influencing the probability of an instructor's performance evaluation ranking.

2.4 Research Framework and Hypotheses Development

Given the objectives of the study and past literature, a research framework is developed as shown in Fig. 1. While the evaluation of instructors' performance is the outcome variable, the three psychological factors (motivation, degree of confidence and level of attention) serve as independent variables. Teaching abilities and physical attractiveness are constructed as antecedents. All hypotheses postulate positive relationships between the variables. Additionally, this study further examines motivation, degree of confidence, and level of attention as mediators on the relationship between antecedents (teaching abilities and physical attractiveness) and performance evaluation.



H1a: *Instructors' teaching abilities have positive effect on students' motivation.*

H1b: *Instructors' teaching abilities have positive effect on students' degree of confidence.*

H1c: *Instructors' teaching abilities have positive effect on students' level of attention.*

H2a: *Instructors' physical attractiveness has positive effect on students' motivation.*

H2b: *Instructors' physical attractiveness has positive effect on students' degree of confidence.*

H2c: *Instructors' physical attractiveness has positive effect on students' level of attention.*

H3: *Students' motivation has positive effect on their evaluation of instructors' performance.*

H4: *Students' degree of confidence has positive effect on their evaluation of instructors' performance.*

H5: *Students' level of attention has positive effect on their evaluation of instructors' performance.*

H6a: *Degree of confidence mediates the relationship between physical attractiveness and evaluation of instructors' performance.*

H6b: *Degree of confidence mediates the relationship between teaching abilities and evaluation of instructors' performance.*

H7a: *Level of attention mediates the relationship between physical attractiveness and evaluation of instructors' performance.*

H7b: *Level of attention mediates the relationship between teaching abilities and evaluation of instructors' performance.*

H8a: *Students' motivation mediates the relationship between physical attractiveness and evaluation of instructors' performance.*

H8b: *Students' motivation mediates the relationship between teaching abilities and evaluation of instructors' performance.*

Fig. 1 Research framework

3 Methodology

The study adopted causal design since the purpose was to examine the cause-and-effect relationships between the variables of interest (Zikmund et al. 2013). It was a cross-sectional study and a quantitative approach using questionnaire survey was used to administer data collection. It was deemed the most appropriate method to test and explain the postulated relationships of the study.

As the university students in Malaysia constituted the target population, a public university was selected as the research site. The reason for this is that every public university

is made up by students from all over the country. Non-probability sampling strategy was employed because there was no way to sample every student on campus with equal chance (Zikmund et al. 2013). As such, both convenience sampling and purposive sampling techniques were utilized to sample the students with the emphasis on the latter to ensure that students sampled were Malaysian undergraduate students. In terms of sample size, 30 to 500 students were considered adequate to conduct a quantitative research (Sekaran and Bougie 2010).

The questionnaire consisted of questions related to student profile, and statements pertaining to the variables under investigation, namely teaching abilities, physical attractiveness, motivation, degree of confidence, level of attention and performance evaluation. 5-point likert scale and multiple items were used for all the statements. A pilot study was carried out to ensure the usability of the questionnaire and the reliability of the measurement. Once the questionnaire was finalized, 200 copies were distributed on campus, and 150 were subsequently collected, yielding a response rate of 75 percent. However, 10 copies were discarded due to obvious straight-lining pattern in the responses and serious data omission.

Data entry was done in Statistical Package for Social Sciences (SPSS) 22.0 but analysis was carried out in SmartPLS 2.0. SmartPLS 2.0 software adopts a variance-based approach, which is also known as partial least squares-structural equation modeling (PLS-SEM). PLS technique was used in this study mainly because of its favourable convergence properties, where it uses separate ordinary least squares (OLS) regressions to estimate the model's partial regression relationships (Ciavolino 2012; Ciavolino and Nitti 2013; Nitti and Ciavolino 2014). In addition, PLS-SEM algorithm assigns more than one latent variable to a block of indicators and imposes orthogonality restrictions among constructs in the structural model (Ciavolino 2012; Nitti and Ciavolino 2014). Hence, PLS-SEM technique is able to test the relationships of the variables in a single measurement and structural models (Hair et al. 2014). Furthermore, the algorithm in PLS-SEM allows a prediction modeling perspective where the method aims to maximize the amount of explained variance of the endogenous latent variables (Hair et al. 2014). As such, PLS-SEM is more suited in the present study as it facilitated the testing of measurement and structural models once and it helped explain the predictive quality of the antecedents and independent variables in the framework.

4 Results and Discussions

4.1 Respondent Profile

A total of 140 students are sampled for the study. Table 1 shows the profile of the students.

4.2 Assessment of Measurement Model

The measurement model is tested to assess the convergent validity and discriminant validity of the reflective constructs (i.e., Level of Attention, Degree of Confidence, Motivation, Physical Attractiveness, Performance Evaluation, and Teaching Abilities) (Hair et al. 2014; Ramayah et al. 2018). Specifically, loading, average variance extracted (AVE), composite reliability (CR), and discriminant validity using cross-loading and Fornell and Lacrker's criterion (1981) are looked at (Hair et al. 2014; Ramayah et al. 2018).

Table 1 Respondent profile

Variables	Categories	Frequency	Percent
Gender	Male	51	36.4
	Female	89	63.6
Races	Malay	52	37.1
	Chinese	59	42.1
	Indian	7	5.0
	Iban	6	4.3
	Others	16	11.4
Program	International Economics	9	6.4
	Industrial Economics	7	5.0
	Service Economics	5	3.6
	Business Economics	9	6.4
	Marketing	36	25.7
	Corporate Management	30	21.4
	Finance	19	13.6
	Accountancy	25	17.9
Year of study	1st Year	50	35.7
	2nd Year	48	34.3
	3rd Year	40	28.6
	4th Year	2	1.4

4.3 Convergent Validity

Table 2 presents the findings of convergent validity test. All AVEs are found to be larger than 0.5, which is suggested in literature to ensure that at least 50 percent of variance is explained by the items of the construct (Hair et al. 2014). Moreover, all CRs are also found to be larger than the threshold of 0.7, indicating that the reliability or consistency of data between the items of the same construct is established (Hair et al. 2014). As such, it is concluded that convergent validity is secured.

4.4 Discriminant Validity

Table 3 represents the cross-loading result. It illustrates that all the indicators' outer loadings are loaded high on the respective constructs and are largest compared to loadings across the constructs. Thus, discriminant validity using cross-loading is achieved.

Table 4 shows the results of discriminant validity with a more stringent test. It is suggested that the square root of each AVE has to be greater than the values of latent variable correlations in order to establish discriminant validity of the data (Vinzi et al. 2010). The findings show that the values on the diagonal are greater than the values beneath and beside. Hence, it is concluded that there is no issue of multi-collinearity among the items of all constructs.

Table 2 Convergent validity results

Construct	Items	Loading	CR[a]	AVE[b]
Level of attention	LA_1	0.770	0.853	0.600
	LA_2	0.517		
	LA_3	0.883		
	LA_4	0.873		
Degree of confidence	DC_1	0.902	0.921	0.795
	DC_3	0.884		
	DC_4	0.890		
Motivation	MO_1	0.834	0.886	0.662
	MO_2	0.673		
	MO_3	0.837		
	MO_4	0.894		
Physical attractiveness	PA_1	0.651	0.851	0.589
	PA_2	0.799		
	PA_3	0.787		
	PA_4	0.823		
Performance evaluation	PE_1	0.808	0.878	0.643
	PE_2	0.788		
	PE_3	0.748		
	PE_4	0.860		
Teaching abilities	TA_1	0.859	0.929	0.767
	TA_2	0.923		
	TA_3	0.890		
	TA_4	0.830		

[a]Composite reliability (CR) = (square of the summation of the factor loadings)/{(square of the summation of the factor loadings) + (square of the summation of the error variances)}

[b]Average variance extracted (AVE) = (summation of the square of the factor loadings)/{(summation of the square of the factor loadings) + (summation of the error variances)}

4.5 Assessment of Structural Model

Structural model is where the hypotheses are tested. Bootstrapping is conducted to normalize the standard errors, thus avoiding deflation or inflation values due to non-normality of data (Wong 2013). The hypotheses are supported when there is no zero straddled in between the lower bound and upper bound of bias corrected accelerated (BCa) confidence interval. Hence, the results as shown in Table 5 demonstrate that the hypotheses of all direct relationships are supported. Specifically, teaching abilities are positively related to motivation (H1a: $\beta = 0.576$, $t = 7.653$), degree of confidence (H1b: $\beta = 0.598$, $t = 9.297$) and level of attention (H1c: $\beta = 0.556$. $t = 7.439$). It corresponds to past literature in general about the importance of instructors' teaching skills to students. Physical attractiveness, in turn, are also found to be positively associated with motivation (H2a: $\beta = 0.167$, $t = 2.216$), degree of confidence (H2b: $\beta = 0.271$, $t = 4.138$) and level of attention (H2c: $\beta = 0.235$, $t = 3.054$), implying the relevance of the instructors' physical attributes and the manner they conduct themselves before the students. Besides, as hypothesized, motivation (H3: $\beta = 0.135$, $t = 1.697$), degree of confidence (H4: $\beta = 0.475$, $t = 3.991$) and level of attention (H5: $\beta = 0.257$, $t = 2.339$) have positive effect on performance evaluation. Hence, how the students perceive the instructors psychologically has an effect on how they evaluate them.

Table 3 Loading and cross loading results

Constructs	LA	DC	MO	PA	PE	TA
LA_1	**0.770**	0.526	0.499	0.326	0.566	0.546
LA_2	**0.517**	0.241	0.354	0.216	0.360	0.197
LA_3	**0.883**	0.706	0.524	0.493	0.623	0.594
LA_4	**0.873**	0.680	0.619	0.512	0.609	0.644
DC_1	0.665	**0.902**	0.679	0.545	0.676	0.734
DC_3	0.697	**0.884**	0.562	0.564	0.685	0.626
DC_4	0.598	**0.890**	0.557	0.441	0.631	0.668
MO_1	0.541	0.664	**0.834**	0.458	0.553	0.665
MO_2	0.358	0.281	**0.673**	0.124	0.330	0.290
MO_3	0.553	0.538	**0.837**	0.382	0.548	0.528
MO_4	0.628	0.608	**0.894**	0.442	0.556	0.577
PA_1	0.352	0.343	0.318	**0.651**	0.282	0.307
PA_2	0.391	0.390	0.330	**0.799**	0.351	0.351
PA_3	0.382	0.487	0.434	**0.787**	0.438	0.472
PA_4	0.466	0.536	0.335	**0.823**	0.499	0.431
PE_1	0.519	0.600	0.475	0.426	**0.808**	0.538
PE_2	0.588	0.530	0.447	0.365	**0.788**	0.404
PE_3	0.517	0.524	0.519	0.362	**0.748**	0.425
PE_4	0.643	0.714	0.567	0.502	**0.860**	0.700
TA_1	0.576	0.624	0.501	0.363	0.522	**0.859**
TA_2	0.617	0.751	0.596	0.506	0.593	**0.923**
TA_3	0.650	0.613	0.631	0.467	0.564	**0.890**
TA_4	0.526	0.668	0.585	0.458	0.619	**0.830**

Bold values are loading values, all exceed than the recommended value of 0.5

LA level of attention, *DC* degree of confidence, *MO* motivation, *PA* physical attractiveness, *PE* performance evaluation, *TA* teaching abilities

Table 4 Discriminant validity results

Constructs	LA	DC	MO	PA	PE	TA
LA	**0.775**					
DC	0.734	**0.892**				
MO	0.654	0.674	**0.814**			
PA	0.521	0.581	0.464	**0.768**		
PE	0.710	0.745	0.628	0.521	**0.802**	
TA	0.677	0.759	0.662	0.515	0.656	**0.876**

Diagonals represent the square root of the AVE (shown as bold) while the other entries represent the correlations of each construct with other constructs

Additionally, the effect size (f^2) is also measured in this study. Effect size (f^2) is an assessment which looks into substantive significance. It is a measure used to assess the relative impact of a predictor construct on an endogenous construct (Hair et al. 2014). To measure the effect size, Cohen's (1988) benchmark is appropriated, whereby 0.02, 0.15, and 0.35 represent small, medium, and large effects respectively (Cohen 1988). As illustrated in Table 5, while the relationship between teaching abilities and the three psychological factors exhibits a large effect size (H1a: $f^2 = 0.450$; H1b: $f^2 = 0.655$; H1c: $f^2 = 0.454$),

Table 5 Bootstrap results of direct hypotheses

Relationship	Beta	T-value	BCa confidence interval		Effect size (f^2)
			LB	UB	
Teaching ability → Motivation	0.576	7.653**	0.444	0.69	0.450
Teaching ability → Degree of confidence	0.598	9.297**	0.491	0.699	0.655
Teaching ability → Level of attention	0.556	7.439**	0.428	0.672	0.454
Physical attractiveness → Motivation	0.167	2.216**	0.041	0.288	0.038
Physical attractiveness → Degree of confidence	0.271	4.138**	0.158	0.373	0.134
Physical attractiveness → Level of attention	0.235	3.054**	0.107	0.355	0.081
Motivation → Performance evaluation	0.135	1.697*	0.001	0.262	0.024
Degree of confidence → Performance evaluation	0.475	3.991**	0.279	0.670	0.217
Level of attention → Performance evaluation	0.257	2.339**	0.061	0.429	0.068

*$p < 0.05$; **$p < 0.01$ (one-tailed)

a small effect size is observed in the relationship between physical attractiveness and the three psychological factors (H2a: $f^2 = 0.038$; H2b: $f^2 = 0.134$; H2c: $f^2 = 0.081$). Notwithstanding small effect in the latter, the impact of instructors' physical attractiveness on students cannot be ignored (Chin et al. 2003). As for the effect of degree of confidence on performance evaluation, the result shows a medium effect size (H4: $f^2 = 0.217$). Lastly, the effect of motivation and level of attention on performance evaluation is substantively small (H3: $f^2 = 0.024$; H5: $f^2 = 0.068$). This shows the relative importance of degree of confidence among the psychological factors to students' evaluation of instructors' performance.

Table 6 shows R^2 of all endogenous variables, and the results indicate that the variance of each construct is well explained by their respective exogenous variables ($R^2_{\text{Level of Attention}} = 0.499$; $R^2_{\text{Degree of Confidence}} = 0.598$; $R^2_{\text{Motivation}} = 0.458$; $R^2_{\text{Performance Evaluation}} = 0.631$). By using blindfolding procedure (Hair et al. 2014), the result suggests the model has predictive relevance ($Q^2_{\text{Level of Attention}} = 0.113$; $Q^2_{\text{Degree of Confidence}} = 0.185$; $Q^2_{\text{Motivation}} = 0.080$; $Q^2_{\text{Performance Evaluation}} = 0.181$). These results indicate the quality of the structural model in explaining the phenomenon under investigation.

Table 7 shows the results of all the mediated relationships. In order to examine the mediation effect, Preacher and Hayes' (2008) method using bootstrapping was employed because the method does not require distributional assumption of an indirect effect estimation and it is robust when applied to either large or small sample size. The result exhibits that three mediation hypotheses (H6a, H6b, and H7b) are supported because there is no zero straddled in between the lower and upper bounds of BCa confidence interval.

Table 6 R^2 and redundancy results

Construct	R^2	Q^2
Level of attention	0.499	0.113
Degree of confidence	0.598	0.185
Motivation	0.458	0.080
Performance evaluation	0.631	0.181

Table 7 Bootstrap results of mediation hypotheses

Relationship	Beta	T-value	BCa confidence interval	
			LB	UB
Physical attractiveness → Degree of confidence → Performance evaluation	0.129	2.895**	0.055	0.239
Teaching abilities → Degree of confidence → Performance evaluation	0.284	3.621**	0.144	0.456
Physical attractiveness → Level of attention → Performance evaluation	0.060	1.742	−0.009	0.144
Teaching abilities → Level of attention → Performance evaluation	0.143	2.201**	0.026	0.278
Physical attractiveness → Motivation → Performance evaluation	0.022	1.249	−0.002	0.069
Teaching abilities → Motivation → Performance evaluation	0.078	1.604	−0.017	0.173

*$p < 0.05$; **$p < 0.01$ (two-tailed)

Consequently, H6a, H6b and H7b are supported while the remaining hypotheses (H7a, H8a, and H8b) are not.

Specifically, degree of confidence mediates the relationship between physical attractiveness and performance evaluation (H6a: $\beta = 0.129$, $t = 2.895$) as well as the relationship between teaching abilities and performance evaluation (H6b: $\beta = 0.284$, $t = 3.621$). Corresponding to the relative importance of degree of confidence as discussed earlier, it shows this particular psychological factor is a significant mechanism in explaining how the instructors' physical attractiveness would affect the students' evaluation of performance. In a similar fashion, teaching abilities have a positive and indirect impact on performance evaluation through level of attention (H7b: $\beta = 0.143$, $t = 2.201$). However, the results illustrate that the indirect effect of level of attention on physical attractiveness and performance evaluation (H7a: $\beta = 0.060$, $t = 1.742$) is not supported. Moreover, the indirect effect of motivation also does not mediate the paths between physical attractiveness and performance evaluation (H8a: $\beta = 0.022$, $t = 1.249$) as well as teaching abilities and performance evaluation (H8b: $\beta = 0.078$, $t = 1.604$). It denotes that motivation does not intervene the effect of the two antecedents on the evaluation of students towards instructors' performance. It can also be surmised that motivation and level of attention do not impose any effect on the relationship between physical attractiveness and performance evaluation.

5 Implication and Conclusion

One of the most striking findings in recent education research revolves around the importance of students evaluation towards instructional abilities. The factors influencing students perception and evaluation choices are thus essential to gauge the evaluation mechanism. Students evaluation may ideally reflect on their own personal experience towards the course or subject and other facilitating conditions. The findings from this study shows that while students are in a good position to evaluate some aspects of teaching, they could potentially evaluate their instructors based on a reaction to irrelevant or non-instructional characteristics that may not be available in other instructors as seen in the relationship between physical attractiveness and performance evaluation mediated by degree of confidence. Furthermore, general responses towards overall performance effectiveness are not influenced by student's motivation to learn, but rather on degree of confidence and to a

lesser degree, level of attention. This is supported by both the degree of confidence and level of acceptance being the mediators of the relationship between teaching abilities and performance evaluation.

Whilst it is a common practice for universities in Malaysia to recognize strengths and weaknesses of the instructors, any efforts to enhance teaching performances would need to address the system as a whole. The system includes classroom settings, instructors and students, context of the syllabus, learning activities and also the outcomes (Altbach et al. 2009). Universities have evolved in the era of globalization to provide significant opportunities for a variety of knowledge and ways seen by some as ideal, hence the need to obtain purposeful information from students is pivotal. It is noteworthy to mention that the findings by Muda et al. (2012) show that favourable ratings are given towards instructional abilities that correlate with the level of confidence in the delivery of subjects taught. Nevertheless, Eagly et al. (1991) suggest that people tend to pre-conceive that people with good looks possess various kinds of good traits and such positive impression contributes to confidence. All these hold true in the findings of the present study.

One of the ultimate goals of an instructor is the ability to transform and motivate their students by providing knowledge and skills. In contrast with other representative studies in this area, the findings in the context of Malaysian universities found no evidence of the effect of motivation on students' evaluation of instructors' performance. Similar studies by Liaw and Goh (2003) and Chan et al. (2012) confirm that the outcomes of teaching abilities are not associated with enhancing students' motivation to learn. Furthermore students have different levels of motivation and may be influenced by other underlying factors, which warrant a possibility of future exploration in this regard.

The use of students' evaluation and feedback are widely endorsed by instructors, students and management especially for those responsible for delivering instruction and intellectual guidance. As such, students' critical assessment of their experiences at a higher learning institution may be seen as responses to increased calls for access to provide better curriculum and pedagogy for students with diverse educational needs and aspirations. Albeit the popularity of these evaluations for the purpose of providing a diagnostic feedback to the universities about the instructors' teaching abilities, research has shown that the validity of such measurements can be misleading due to strong preferences towards other traits rather than teaching itself (Ottoboni et al. 2016).

While there appears to be a plethora of many empirical works of mitigating factors influencing instructors' performance evaluation, this study substantially adds to the marketing education literature in several notable ways. This research has substantiated physical attractiveness as a plausible influencing factor for students to assess the critical variable of instructional abilities. Although this phenomena is not included in evaluation forms of any public universities in Malaysia and it could be deemed subjective and immeasurable, the existence of such pattern should serve as a precursor for the University's stakeholders and faculty members on students evaluation process. The significant importance of presetting instructors' physical attractiveness in the questionnaire, including suitable outfit and well-groomed appearance, may project a broad overview of performance evaluation process together with existing elements such attitudes and responsibility.

It is extremely difficult to ensure that students' evaluation of an instructor is a true representation of instructional abilities. The crux of this problem may lie in the students' perception of what constitute good teaching. Whilst many universities continue to use students' feedback as a foundation to improve their objectives, in reality students' perception may be distorted as a result of pedagogic discrimination. This conclusion suggests that whilst physical attractiveness of an instructor may have profound influence a student's level

of confidence and attention as well as evaluation process on teaching abilities, a student's attribute as a learner is governed by multiple factors. Hence, this research provides empirical contribution towards students' evaluation of instructors based on the discrimination of physical attractiveness. Suffice to say, there is no clear indication as to whether an instructor's performance is affected by unobserved productive work by students or a pure discrimination of physical appearance. Therefore, this will provide an insightful guidance and possible avenue to further investigate and develop useful tools for evaluating instructors in the contemporary setting.

To move forward, researchers should consider alternative estimation methods of PLS to gain more insights of studies related to performance evaluation in the educational setting. For instance, future studies could consider the use of qualitative external information to examine the phenomenon in PLS-SEM (Ciavolino et al. 2015). Such approach would complement the study as it characterizes the observations according to the context of study as well as identifies potential heterogeneity of the target population, such as difference of socio-demographic factors in gender and education level. Besides, some new estimators in PLS such as the nonlinear principal component analysis (NPCA), Rasch analysis with the Rating Scale Model (NPCA-RSM) and structural equation model based on generalized maximum entropy (SEM-GME) or Generalized Cross Entropy (SEM-GCE) could be incorporated into the relevant studies (Ciavolino et al. 2015; Ciavolino and Calcagnì 2015; Ciavolino and Carpita 2015). These novel approaches are able to maximize the homogeneity of the trait as well as allow greater reduction of redundancy without sacrificing any measurement information. In other words, they facilitate the reduction of the variables under investigation and the evaluation of measurement errors before assessing the structural relationships of the research model.

Acknowledgements The authors would like to thank the two anonymous referees and the editor for their helpful comments and suggestions on an earlier drafts. The authors gratefully acknowledges financial support from Universiti Malaysia Sarawak (UNIMAS) Geran Penyelidikan Khas (Top Down) 03(TD04)/1054/2013(02). As usual, the responsibility of errors and omissions rests with the authors.

References

Altbach, P. G., Reisberg, L., & Rumbley, L. E. (2009). *Trends in global higher education: Tracking an academic revolution*. Paris: UNESCO Publishing.

Bakar, A. R., Mohamed, S., & Zakaria, N. S. (2013). They are train to teach, but how confience thay are? A study of student teachers' sense of efficiency. *Journal of Social Sciences, 8*(4), 497–504.

Barrie, S., & Ginns, P. (2007). The linking of national teaching performance indicators to improvements in teaching and learning in classrooms. *Quality in Higher Education, 13*(3), 275–286.

Bassi, F., Clerci, R., & Aquario, D. (2017). Students' evaluation of teaching at a large Italian university: Measurement scale validation. *Electronic Journal of Applied Statistical Analysis, 10*(1), 93–117.

Bénabou, R., & Tirole, J. (2002). Self confidence and personal motivation. *The Quarterly Journal of Economics, 117*(3), 871–915.

Benton, S. L., & Cashin, W. E. (2014). Student ratings of instruction in college and university courses. In M. B. Paulsen (Ed.), *Higher education: Handbook of theory and research* (pp. 279–326). Dordrecht: Springer.

Byrne, M., & Flood, B. (2005). A study of accounting students' motives, expectations and preparedness for higher education. *Journal of Further and Higher Education, 29*(2), 111–124.

Chan, B. Y.-F., Yeoh, S. F., & Ho, J. S.-Y. (2012). Student evaluation of lecturer performance among private university students. *Canadian Social Science, 8*(4), 238–243.

Chin, W. W., Marcolin, B. L., & Newsted, P. R. (2003). A partial least squares latent variable modeling approach for measuring interaction effects: Results from a Monte Carlo simulation study and an electronic-mail emotion/adoption study. *Information System Research, 14*(2), 189–217.

Ciavolino, E. (2012). General distress as second order latent variable estimated through PLS-PM approach. *Electronic Journal of Applied Statistical Analysis, 5*(3), 458–464.
Ciavolino, E., & Calcagnì, A. (2015). Generalized cross entropy method for analysing the SERVQUAL model. *Journal of Applied Statistics, 42*(3), 520–534.
Ciavolino, E., & Carpita, M. (2015). The GME estimator for the regression model with a composite indicator as explanatory variable. *Quality & Quantity, 49*(3), 955–965.
Ciavolino, E., Carpita, M., & Nitti, M. (2015). High-order pls path model with qualitative external information. *Quality & Quantity, 49*(4), 1609–1620.
Ciavolino, E., & Nitti, M. (2013). Using the hybrid two-step estimation approach for the identification of second-order latent variable models. *Journal of Applied Statistics, 40*(3), 508–526.
Cipriani, G. P., & Zago, A. (2011). Productivity or discrimination? Beauty and the exams. *Oxford Bulletin of Economics and Statistics, 73*(3), 428–447.
Coates, H. (2005). The value of student engagement for higher education quality assurance. *Quality in Higher Education, 11*(1), 25–36.
Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.
Copland, F., Ma, G., & Mann, S. J. (2009). Reflecting in and on post-observation feedback in initial teacher training on certificate courses. *English Language Teacher Education and Development, 12,* 14–23.
De Paola, M., & Scoppa, V. (2015). Gender discrimination and evaluators' gender: Evidence from Italian academia. *Economica, 82*(325), 162–188.
Dion, K., Berscheid, E., & Walster, E. (1972). What is beautiful is good. *Journal of Personality and Social Psychology, 24*(3), 285.
Eagly, A. H., Ashmore, R. D., Makhijani, M. G., & Longo, L. C. (1991). What is beautiful is good, but…: A meta-analytic review of research on the physical attractiveness stereotype. *Psychological Bulletin, 110*(1), 109.
Eison, J. (1990). Confidence in the classroom: Ten maxims for new teachers. *College Teaching, 38*(1), 21–25.
Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research, 18*(1), 39–50.
Fulmer, S. M., & Turner, J. C. (2014). The perception and implementation of challenging instruction by middle school teachers. *The Elementary School Journal, 114*(3), 303–326.
Gurung, R. A., & Vespia, K. (2007). Looking good, teaching well? Linking liking, looks, and learning. *Teaching of Psychology, 34*(1), 5–10.
Hair, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2014). *A primer on Partial Least Squares Structural Equation Modeling (PLS-SEM)*. Los Angeles, CA: Sage Publications Inc.
Hamermesh, D. S., & Parker, A. (2005). Beauty in the classroom: Instructors' pulchritude and putative pedagogical productivity. *Economics of Education Review, 24*(4), 369–376.
Hatfield, E., & Rapson, R. L. (2000). Physical attractiveness. *The Corsini Encyclopedia of Psychology and Behavioral Science, 3,* 1203–1205.
Ingusci, E., Palma, F., De Giuseppe, M., & Iacca, C. (2016). Social and scholar integration and students satisfaction: The mediating role of career adaptability. *Electronic Journal of Applied Statistical Analysis, 9*(4), 704–715.
Johnson, S. M. (2012). Having it both ways: Building the capacity of individual teachers and their schools. *Harvard Educational Review, 82*(1), 107–122.
Klassen, R. M., & Chiu, M. M. (2010). Effects on teachers' self-efficacy and job satisfaction: Teacher gender, years of experience, and job stress. *Journal of Educational Psychology, 102*(3), 741–756.
Kogan, L. R., Schoenfeld-Tacher, R., & Hellyer, P. W. (2010). Student evaluations of teaching: Perceptions of faculty based on gender, position, and rank. *Teaching in Higher Education, 15*(6), 623–636.
Kurtulmuş, B. E., Warner, B., & Özarı, Ç. (2016). Research or teaching oriented? Game theory models for the strategic decision-making process of universities with the external environment held neutral. *Electronic Journal of Applied Statistical Analysis, 9*(3), 469–490.
Liaw, S. H., & Goh, K. L. (2003). Evidence and control of biases in student evaluations of teaching. *International Journal of Educational Management, 17*(1), 37–43.
MacNell, L., Driscoll, A., & Hunt, A. N. (2015). What's in a name: Exposing gender bias in student ratings of teaching. *Innovative Higher Education, 40*(4), 291–303.
Marsh, H. W. (2007). Students' evaluations of university teaching: Dimensionality, reliability, validity, potential biases and usefulness. In R. P. Perry & J. C. Smart (Eds.), *The scholarship of teaching and learning in higher education: An evidence-based perspective* (pp. 319–383). Dordrecht: Springer.

McGrath, B., Brennan, M., Dolan, P., & Barnett, R. (2009). Adolescent well-being Electronic Journal of Applied Statistical Analysis 715 and supporting contexts: A comparison of adolescents in Ireland and Florida. *Journal of Community & Applied Social Psychology, 19*(4), 299–320.

Mobius, M. M., & Rosenblat, T. S. (2006). Why beauty matters. *American Economic Review, 96*(1), 222–235.

Moss-Racusin, C. A., Dovidio, J. F., Brescoll, V. L., Graham, M. J., & Handelsman, J. (2012). Science faculty's subtle gender biases favor male students. *Proceedings of the National Academy of Sciences, 109*(41), 16474–16479.

Mousavi, A., Mares, C., & Stonham, T. J. (2015). Continuous feedback loop for adaptive teaching and learning process using student surveys. *International Journal of Mechanical Engineering Education, 43*(4), 247–264.

Muda, N., Samsudin, H. B., Majid, N., Ali, K. A. M., & Ismail, W. R. (2012). Students perspective on lecturer characteristic for effective teaching. *Procedia-Social and Behavioral Sciences, 59,* 535–540.

Nitti, M., & Ciavolino, E. (2014). A deflated indicators approach for estimating second-order reflective models through PLS-PM: An empirical illustration. *Journal of Applied Statistics, 41*(10), 2222–2239.

Olson, I. R., & Marshuetz, C. (2005). Facial attractiveness is appraised in a glance. *Emotion, 5*(4), 498–502.

Ottoboni, K., Boring, A., & Stark, P. (2016). Student evaluations of teaching (mostly) do not measure teaching effectiveness. *Science Open Research*. https://doi.org/10.14293/s2199-1006.1.sor-edu.aetbzc.v1.

Ponzo, M., & Scoppa, V. (2013). Professors' beauty, ability, and teaching evaluations in Italy. *The B.E. Journal of Economic Analysis & Policy, 13*(2), 811–835.

Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods, 40*(3), 879–891.

Putman, S. M. (2012). Investigating teacher efficacy: Comparing pre-service and in service teachers with different levels of experience. *Action in Teacher Education, 34*(1), 26–40.

Ramayah, T., Cheah, J., Chuah, F., Ting, H., & Memon, M. A. (2018). *Partial least squares structural equation modeling (PLS-SEM) using SmartPLS 3.0: An updated and practical guide to statistical analysis* (2nd ed.). Singapore: Pearson.

Ryan, R. M., & Deci, E. L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology, 25*(1), 54–67.

Sekaran, U., & Bougie, R. (2010). *Research method for business: A skill building approach*. Chichester, West Sussex: Wiley.

Simonetti, B., Sarnacchiaro, P., & Rosario Gonzalez Rodrıguez, M. (2017). Goodness of fit measures for logistic regression model: An application for students' evaluations of university teaching. *Quality & Quantity, 51*(6), 2545–2554.

Smimou, K., & Dahl, D. W. (2012). On the relationship between students' perceptions of teaching quality, methods of assessment, and satisfaction. *Journal of Education for Business, 87*(1), 22–35.

Spooren, P., Brockx, B., & Mortelmans, D. (2013). On the validity of student evaluation of teaching the state of the art. *Review of Educational Research, 83*(4), 598–642.

Tschannen-Moran, M., & Hoy, A. W. (2001). Teacher efficacy: Capturing an elusive construct. *Teaching and Teacher Education, 17*(7), 783–805.

Vinzi, V., Chin, W. W., Henseler, J., & Wang, H. (2010). *Handbook of partial least squares: Concepts, methods and applications*. New York, NY: Springer.

Wong, K. K. K. (2013). Partial Least Squares Structural Equation Modeling (PLS-SEM) techniques using Smart PLS. *Marketing Bulletin, 24,* 1–32.

Worthington, A. C. (2002). The impact of student perceptions and characteristics on teaching evaluations: A case study in finance education. *Assessment & Evaluation in Higher Education, 27*(1), 49–64.

Zikmund, W. G., Babin, B. J., Carr, J. C., & Griffin, M. (2013). *Business research methods*. Cengage Learning.