Oral Raters' Stress, Motivation and Work Performance

Sofia Tsagdi[1], Dimitrios Karakoulas[1] & Kostas Theologou[1]

[1] National Technical University of Athens, Greece

Correspondence: Sofia Tsagdi, Department of Law and Social Sciences Athens, 1528, National Technical University of Athens, Greece.

Abstract

The aim of this study is to investigate oral raters' stress and motivation and the relation of these factors to work performance. In order to examine this issue thoroughly, we collected statistical data from 94 oral raters working in different English language certifiers in Greece. The quantitative results were also triangulated with the managers' and supervisors' comments, which were excerpted from semi-structured interviews, with the aim of arriving to a more representative picture. Work motivation, commitment and work stress were correlated with social identity factors such as gender, educational level, nationality and work experience. The results of our study concluded that a significant proportion of oral raters reported that low levels of work stress, high motivation and high work commitment can positively affect high quality service and high work performance. We further found that work motivation, task performance and work stress are positively related to personal identification. Implications of these findings are discussed proving to a large extent that a positive work environment and good working conditions can actually guarantee high work performance.

Keywords: oral raters, stress, working conditions, work performance

1. Introduction

1.1 Raters' Perception of the Work Environment

Most studies concerning oral raters have focused on rater variability, rater bias and on uncovering the characteristic differences between rater backgrounds on test scores. For years, oral raters' performance has been examined and researched in terms of rating experience, language background, effectiveness of training (Kim, 2009; Zang & Elder, 2011; Xi & Mollaun, 2009), with yet little discussion about the impact of employees' perceptions of their working environment, motivation, stressors and their potential significant implications for work productivity. Although these factors have indeed been examined, in the last decades, research in various fields and thousands of studies have focused on job satisfaction, organization commitment, work stress and similar topics (Brayfield and Crockett, 1955; Muchinsky, 1985; Schuler, 1980).

1.2 Aim of the Study

The aim of the current study is to commence a thorough re-examination of the aforementioned issues taking into account the oral rater's perspective. Thus, we hope to begin to investigate and re-examine these issues in the context of oral rater's position. Interesting conclusions based on elaborate research in different fields infer that satisfied, motivated employees provide quality services that in turn offer a competitive advantage in many institutions and industries. In a similar vein, a number of researchers have argued for and have produced evidence establishing a link between quality service and measures of job satisfaction (Hoffman and Ingram, 1992; Schlesinger and Zornitsky, 1992; Schneider and Bowen, 1985). Furthermore, work stress will be another crucial variable examined in our research. Although stress is a highly researched topic, studies examining actual job performance in relation with work stress are fewer. Many studies investigate the relationship between stress and physiological or attitudinal changes (Anderson, 1991; Ganster et al., 1986; Lundberg et al., 1989; Meglino, 1977), but not the dynamic dipole: stress and performance. Our study aims at addressing this gap in research. Moreover, we wish to parallel those researches and further focus on job stress, organization structure, motivation and performance.

In particular, our goal is to examine how perceptions of three job stressors, namely work load, role overload, role conflict and role ambiguity, potentially impact on the effectiveness of services delivered by oral raters. Several researches on service jobs have pointed out the existence of a classic boundary-spanning position, one where the

incumbent might feel pulled in several directions at the same time (Montgomery et al., 1996; Singh et al., 1994; Varca, 1992).

1.3 Research Design

As a first step in the project, the researchers conducted structured one-on-one interviews with oral examiners service consultants and a number of oral raters and proctors. The interviews revealed a common theme with respect to the work environment: participants mentioned stress as a job characteristic. They discussed having a lack of control over their jobs, an inability to satisfy conflicting demands from "clients" and organization's personnel, and a sense of being overburdened. In total, they were offering a description of what the literature would characterize as a stressful work environment, increased role conflict, role overload and role ambiguity. Based on their comments and on relevant literature, we designed and piloted our online questionnaires. In particular, job stress (Independent variable) was measured with the Kahn et al. (1964) instrument which employs a 15-item Likert scale format. The Kahn et al. (1964) scale used here fits within this environmental school; it measures perceptions of the work context. Admittedly, this scale represents only one strand in the complex nomenological net referred to as job stress. However, it operationalizes the stress construct in terms almost identical to the findings reported in initial interviews with incumbents and, therefore, seemed appropriate. Motivation was measured using Perry's (1996) motivation scale for public service motivation. Organizational Commitment Questionnaire (OCQ) was used to measure employee's organizational commitment. It is a 15-item scale developed by Mowday, Steers and Porter (Mowday et al., 1979) and uses a 5-point Likert type response format, with 3 factors that can describe this commitment: acceptance of organizational values. Our dependent variable is job performance. Initially, from a list of 42 generic abilities, a sample of service consultants, organizers and supervisors ranked the top ten abilities (i.e. oral communication) required for job success. These results were subsequently discussed in separate focus groups with service supervisors (three/four per group). The purpose was to finalize a performance evaluation form that would be used as the dependent variable. Seven performance factors were ultimately selected as critical to excellent raters' service:

Coordination

Interpersonal skills

Record Keeping

Problem Solving

Accurate Grading

Inter-rater collaboration

Written feedback

2. Method

The sample of our study includes 94 participants (N=94), all of whom have been certified by their organization as English Language Oral Raters. The main instrument used for data collection is a computer-based questionnaire that has been forwarded via email to the participants through a single focal point, namely the Coordinator of the exam organization. Throughout this process, it has been stressed that the participation is optional and on a volunteering basis, with no effect on the relationship between employer and employees. Furthermore, the participants have been assured that all or part of their responses provided throughout this study will not be communicated by any means to their employer, in order to eliminate the chances of a possible breach of confidentiality. This manipulation has been considered vital, on the one hand in order to address biased errors due to the working relationship between the employer and the employees and on the other hand in order to protect the participants from coercion or undue influence as it has been concluded by Resnik (Resnik, 2016).

Three main scales have been included in the aforementioned instrument, namely the scale that measures commitment to Monday, Steers and Porter scale, the Perry scale that measures participant's motivation to work performance, and Kahn scale that measures stress levels of the participants. Education level, gender and place of education data have also been included in the questionnaire. The main hypothesis under test is that fact that the commitment to the organization of the English Language oral raters and their motivation to provide their services have a negative correlation with their stress level during the examination procedure. A secondary hypothesis that needs to be tested is if a third parameter, namely the education level of the participants, has also a negative correlation with the stress level.

The Monday, Steers and Porter commitment Scale consists of 6 Items, the Perry motivation scale consists of 8 items and the Kahn stress level scale consists of 9 items, all of which are shown in Table 1.

Table 1. Motivation, Commitment and Stress Level Scales (Items Description)

Item Code	Variable	Item Description
V1_1		I feel a strong sense of commitment to my organization
V1_2		I feel like a part of the family at my organization
V1_3	Commitment	I feel emotionally attached to my organization
V1_4		My personal values match my organization's values and culture
V1_5		My pedagogical values in life are very similar to the things my organization's values
V1_6		My organizations values and culture provide a good fit with the things I value in life
V2_1		I find real satisfaction in my work
V2_2		I find real enjoyment in my work
V2_3		I like my job but I consider it more as a way to make an income
V2_4		I am satisfied with the payment I receive for my services
V2_5	Motivation	I am satisfied with the method of payment for my services
V2_6		How much do you really want to make a good work?
V2_7		How much do you feel your personal satisfactions are related to how well you do your job?
V2_8		I look forward with enjoyment to work
V3_1		I feel bad when I make mistakes in my work
V3_2		Do you feel supported If you make a mistake in your wok?
V3_3		Do you seek assistance when you are found in difficult situation?
V3_4		Do you show and report that there are things that are bothering you?
V3_5	Stress Level	Do you seek sympathy and comfort from your supervisors when in a difficult situation?
V3_6		When you made mistake at work did you find it difficult to relax?
V3_7		I feel the training I received well prepared me for the actual work in the field
V3_8		I feel that the training I received matches the demands of the actual work
V3_9		I feel that the training has well prepared me for the evaluation I receive on the job

A 7 point Likert scale pattern is utilized for all three variables (1 Strongly Disagree to 7 Strongly Agree). All items have received a unique item code to be used hereinafter. The first step, once data collection phase has been completed, is to build a reliable model, in order to test the hypotheses. To this end, dimension reduction has been performed for all 3 scales with the use of SPSS 26. The main goal is to reduce the number of items per scale towards reaching a more parsimonious model. Reliability analysis and Exploratory Factor Analysis (EFA) have been performed.

More specifically, reliability analysis for all three scales aimed at Cronbach's alpha > .70 (Taber, 2017), allowing the option to exclude items in order to achieve the desired reliability level. EFA has been performed to detect the specific factors extracted from the sample of this study, including principal axis factoring extraction and Oblimin rotation methods. The number of factors to be extracted are chosen as per the scree plot (Cattell, 1966) and the Kaiser's criterion (eigenvalues > 1). Items were retained when rotated loadings were close or greater than .30 (Stevens, 2002). The following EFA assumptions have been taken into account: a) no significant outliers in the items, examined with boxplots, b) sample adequacy as per the Kaiser-Meyer-Olkin index (KMO) greater than 0.5 (Field, 2017) and c) significant linear correlations between the items of about .3 or greater. Some items have been reversed, since the direction of responses is required to be common, in order to perform meaningful statistical analysis. The final utilized scales have proven reliability and validity and the remaining items per scale provide a more accurate set for averaging.

The next step that follows the production of the single dimensional model for the 3 variables, is to perform regression analysis, in order to compare three independent variables, namely the motivation, commitment and the education level with one dependent variable, namely the stress level. The education level is measured through a categorical scale, however, in order to be included in the regression analysis, it has been transformed to a 4 point Likert scale following the suggestions of Gaito (Gaito, 1980). Linear Regression has been performed

with stepping method criteria with probability F of entry .05 and removal .10. Combined interaction of 3 variables have also been considered (7 discrete combinations) and statistical significance p-value < .05 has been set as the acceptable criterion throughout this study (Cumming, 2012).

(1)

3. Results

Descriptive statistical analysis has shown that, the majority of our sample (89.4%) consists of female examiners, who were educated mainly in Greece (64.9%). On the other hand, the distribution of the education level is more evenly distributed across the selected categories. More specifically, Secondary education scored 33.3%, Bachelor degree scored 31.1%, 33.3% hold a Master degree and 2.2% hold a PhD degree. Reliability analysis provided acceptable results with all three scales reaching Cronbach's alpha > .7 as shown in Table 2.

Table 2. Reliability Tests Results

Variable	Cronbach's Alpha	Cronbach's Alpha (after exclusion of items)
Commitment	0.885	0.885
Motivation	0.605	0.804
Stress	0.779	0.779

One item from motivation scaled (V2_3) had to excluded from the final model, in order to achieve the acceptable specification. Measurements received with the commitment achieved a high reliability score reaching almost .89.

As far as the EFA is concerned, the sample adequacy has been proven through the KMO and Bartlett's test performed with SPSS.

Table 3. Kaiser-Meyer-Olkin index (KMO) and Bartlett's Test Results

Variable	Kaiser-Meyer-Olkin Measure of Sampling Adequacy
Commitment	.833
Motivation	.645
Stress	.713

Table 3 summarizes the results for all three scales. Structure Matrices for Commitment Motivation and Stress Level are shown in Tables 4, 5 and 6 respectively.

Table 4. Structure Matrix for Commitment

Item Code	Component 1
V1_1	.695
V1_2	.780
V1_3	.801
V1_4	.878
V1_5	.843
V1_6	.858

Table 5. Structure Matrix for Motivation

Item Code	Component 1	Component 2	Component 3
V2_1	.914		
V2_2	.848		
V2_4		.872	
V2_5		.893	
V2_6			.876
V2_7	.466		.721
V2_8	.810		

Table 6. Structure Matrix for Stress Level

Item Code	Component 1	Component 2	Component 3
V3_1			.959
V3_2	.443	.748	.477
V3_3	.430	.739	.498
V3_4		.744	
V3_5		.693	
V3_7	.952		
V3_8	.958		
V3_9	.952		

EFA showed a 3 factor structure for Stress Level and Motivation and a 1 factor model for the Commitment variable. Hence, direct oblimin rotation was finally performed only for the former two variables. Items with loading greater than .80 were those that were included in the final model.

Table 7. Dimension Reduction Process

Variable	Initial Number of Items per Variable	Number of Items after Reliability Test	Final Number of Items after EFA	Items Selected (Code)
Commitment	6	6	4	V1_3 V1_4 V1_5 V1_6
Motivation	8	7	6	V2_1 V2_2 V2_4 V2_5 V2_6 V2_8
Stress	9	8	4	V3_1 V3_7 V3_8 V3_9

Table 7 summarizes the dimension reduction process, including both EFA and reliability analysis.

Table 8. Final Items of the Model

Variable	Item Code
Commitment	V1_3 V1_4 V1_5 V1_6
Motivation	V2_1 V2_2 V2_4 V2_5 V2_6 V2_8
Stress	V3_1 V3_7 V3_8 V3_9

Table 8 shows the final items of our model that will be averaged to produce single dimensional variables. Finally, regression analysis results, including combined interaction, are summarized in Table 9.

Table 9. Regression Analysis Results

	Variable/Interaction of Variables	Correlation with Stress Variable
Pearson Correlation	Commitment (C)	-.423
	Motivation (M)	-.412
	Education Level (EL)	-.126
	EL X C X M	.057
	EL X M	-.351
	M X C	.019
	EL X C	.075
Sig. (1-tailed)*	Commitment (C)	.000
	Motivation (M)	.000
	Educational Level (EL)	.113
	EL X C X M	.292
	EL X M	.000
	M X C	.429
	EL X C	.236

X Signifies interaction between variables; * Stat. Significance α < .05

The primary test hypothesis has been verified, since both commitment and motivation have demonstrated a significant negative correlation (.-423 and -.412 respectively) as per the regression analysis that has been performed. The combined interaction of motivation and commitment did not produce statistically significant results (p-value > .05). This finding shows that either of the two factors suffice to produce lower stress levels. Classification of participants in the education level category produced more evenly distributed results, with respect to gender and place of education, as mentioned earlier. Therefore, it has been selected as the third variable in the secondary hypothesis. However, only partial verification of this hypothesis has been achieved, since only the combined interaction of education level and motivation produced statistically significant results and showed a negative correlation of -351.

4. Discussion

Our study is in line with the on-growing body of literature that supports that stress, lack of motivation and work commitment cause work performance to suffer and negatively affect service quality. Our results add to the already available literature that the oral raters' educational level and their genre are key-factors that should be seriously taken into consideration in terms of work performance. What is surprising in our results is the fact that, although oral examiners were under strict control and had little "freedom" to underproduce, a low –yet worth mentioning– percentage of participants felt stress over the quality of their work and the quality of the results they provided. At the same time, they claimed that the lack of initiative and creative approaches to problem solving further increased stress and weakened their productivity. On the contrary, raters who felt low work stress engaged in their work more eagerly and effectively.

One straightforward implication and recommendation, which can certainly be drawn from our study, is that work stressors, low motivation and commitment can lower work quality and service against an organization's goals. What is more, conducting in-house surveys in these fields regularly is essential to ensure vigilance for potential problems. also highly recommended. Although most managers may claim that all jobs are stressful, having direct data from employees handing day-to-day fires fights is crucial for securing a consensus about interventions.

We feel that our research further showed that, regardless of the strict controls and validation processes, this particular job should also be studied within the broader organizational context. Otherwise, if the pace of work is hectic, not only the control spam for supervisors is too broad to handle, but also individual work performance may suffer and that can be difficult to be overseen. Finally, our findings suggest that skilled examiners are willing to provide quality services when they feel that their views and work environment are positive.

The message, we believe, is clear: "spend less time supervising and more time creating and developing a work environment that is a good place to work". Our study further validated the current claim that organizations need to rely more on highly motivated and committed employees rather than in conventional organizational controls, so as to ensure high work performance results.

References

Brayfield, A. H., & Crockett, W. H. (1955). Employee attitudes and employee performance. *Psychological bulletin, 52*(5), 396-424. https://doi.org/10.1037/h0045899

Cattell, R. (1966). The Scree Test For The Number Of Factors. *Multivariate Behavioral Research, 1*(2), 245-276. https://doi.org/10.1207/s15327906mbr0102_10

Cumming, G. (2011). *Understanding The New Statistics: Effect Sizes, Confidence Intervals, and Meta-Analysis* (1st ed.). New York, USA: Routledge. https://doi.org/10.4324/9780203807002

Field, A. (2017). *Discovering statistics using IBM SPSS statistics* (5th ed.). Los Angeles, USA: SAGE.

Gaito, J. (1980). Measurement scales and statistics: Resurgence of an old misconception. *Psychological Bulletin, 87*(3), 564-567. https://doi.org/10.1037/0033-2909.87.3.564

Hoffman, K. D., & Ingram, T. N. (1992). Service provider job satisfaction and customer. *Journal of Services Marketing, 6*(2), 68-78. https://doi.org/10.1108/08876049210035872

Iaffaldano, M. T., & Muchinsky, P. M. (1985). Job satisfaction and job performance: A meta-analysis. *Psychological bulletin, 97*(2), 251-273. https://doi.org/10.1037/0033-2909.97.2.251

Kahn, R. L., Wolfe, D. M., Quinn, R. P., Snoek, J. D., & Rosenthal, R. A. (1964). *Organizational stress: Studies in role conflict and ambiguity*. USA: John Wiley.

Kim, A. Y., & Gennaro, K. D. (2012). *Scoring behavior of native vs. non-native speaker raters of writing exams*. Retrieved from https://s-space.snu.ac.kr/bitstream/10371/86486/1/5.%202223404.pdf

Mowday, R. T., Steers, R. M., & Porter, L. W. (1979). The measurement of organizational commitment. *Journal of vocational behavior, 14*(2), 224-247. https://doi.org/10.1016/0001-8791(79)90072-1

Perry, J. L. (1996). Measuring public service motivation: An assessment of construct reliability and validity. *Journal of public administration research and theory, 6*(1), 5-22. https://doi.org/10.1093/oxfordjournals.jpart.a024303

Resnik, D. B. (2016). Employees as Research Participants: Ethical and Policy Issues. *IRB: Ethics and Human Research, 38*(4), 11-16. Retrieved from https://www.researchgate.net/profile/David-Resnik-2/publication/306978670_Employees_as_research_participants_Ethical_and_policy_issues/links/5cbf05dea6fdcc1d49a8c73f/Employees-as-research-participants-Ethical-and-policy-issues.pdf

Schneider, B., & Bowen, D. E. (1985). Employee and customer perceptions of service in banks: Replication and extension. *Journal of applied Psychology, 70*(3), 423-433. https://doi.org/10.1037/0021-9010.70.3.423

Schuler, R. S. (1980). Definition and conceptualization of stress in organizations. *Organizational behavior and human performance, 25*(2), 184-215. https://doi.org/10.1016/0030-5073(80)90063-X

Singh, J., Goolsby, J. R., & Rhoads, G. K. (1994). Behavioral and psychological consequences of boundary spanning burnout for customer service representatives. *Journal of Marketing Research, 31*(4), 558-569. https://doi.org/10.1177/002224379403100409

Stevens, J. (2002). *Applied multivariate statistics for the social sciences* (4th ed.). Hillsdale, NJ: Erlbaum.

Taber, K. S. (2018). The Use of Cronbach's Alpha When Developing and Reporting Research Instruments in Science Education. *Research in Science Education, 48*(6), 1273-1296. https://doi.org/10.1007/s11165-016-9602-2

Varca, P. E. (1992). Work stress and customer service delivery. *Journal of Services Marketing, 13*(3), 229-241. https://doi.org/10.1108/08876049910273853

Xi, X., & Mollaun, P. (2009). *How do raters from India perform in scoring the TOEFL iBT speaking section and what kind of training helps?* (TOEFL iBT Research Report No. RR-09-31). Princeton, NJ: ETS. https://doi.org/10.1002/j.2333-8504.2009.tb02188.x

Zhang, Y., & Elder, C. (2011). Judgments of oral proficiency by non-native and native English speaking teacher raters: Competing or complementary constructs? *Language Testing, 28*(1), 31-50. https://doi.org/10.1177/0265532209360671