

Self-Ratings of Workplace Behaviour: Contrasting Russia and Poland with the United States

Robert D. Costigan

St. John Fisher College

Richard C. Insinga

State University of New York, Oneonta

Grazyna Kranas

University of Warsaw, Poland

Selim S. Ilter

St. John Fisher College

J. Jason Berman

St. John Fisher College

Vladimir A. Kureshov

Siberian State Aerospace University, Russia

This study examines the self-rating appraisal process, assessing the ratings made by the subordinate and the supervisor. Others have proposed that employees in collectivist countries tend to give harsher ratings of themselves than their supervisors, based on data from Taiwan. Conflicting results from more recent Chinese studies have cast doubt on this hypothesis. The present study tests this proposition on data collected in the U.S. and two other collectivist countries, Poland and Russia. ANCOVA results indicate that, as expected, the Polish self-ratings show a harshness bias, but the Russian findings tend to show a leniency effect and the U.S. results tend to show a harshness bias.

Human resource (HR) scholars (e.g., Elenkov, 1998) have advocated for the adoption of performance appraisals in Eastern European firms. Knowing more about where and how different appraisal practices are effective will help HR practitioners in these firms to develop appropriate appraisal systems. This study addresses one aspect of Western-style performance appraisals (i.e., employee self-ratings) using data from Russian, Polish, and U.S. firms.

Ratings completed on one's own performance, otherwise known as self-ratings, are used to improve communication between supervisors and employees in the appraisal meeting. Prior research, however, indicates that self-ratings and the boss' ratings can be expected to disagree. Biased self-ratings has been reported in the U.S. self-rating literature (Harris & Schaubroeck, 1988) and also in international studies done in Taiwan (Farh, Dobbins, & Cheng, 1991) and China (Yu & Murphy, 1993). Therefore, it is necessary to understand the presence, magnitude, and direction of bias when using such ratings in other countries, such as Poland and Russia.

Self-Rating Research in the International Setting

Three notable studies (Farh et al., 1991; Furnham & Stringfield, 1994; Yu & Murphy,

1993) examined self-ratings and corresponding supervisory ratings in a cross-cultural context. Farh et al. (1991) found that Taiwanese employees rated their own performance more harshly than did their bosses. That is, a modesty bias in these Taiwanese self-ratings was detected when the ratings were compared with the supervisor's ratings of the worker. Farh et al. (1991) attributed this modesty bias in self-ratings to collectivist values rooted in the Taiwanese culture. Because collectivist cultures emphasize harmony in relationships (e.g., Hofstede, 2001), there is pressure for workers to understate individual accomplishments and exhibit personal modesty, thus yielding lower-than-expected self-ratings. Providing added support for this, Markus and Kitayama (1991) stated that persons in interdependent cultures tend to be oriented to others instead of the self, and such self-other construal may yield a habitual modest-response tendency.

In an attempt to replicate Farh et al.'s (1991) findings in another collectivist culture, mainland China, Yu and Murphy (1993) found self-ratings to be significantly higher for the Chinese employees when compared to their supervisors' ratings. The modesty bias found in Farh et al.'s (1991) Taiwan study did not surface in Yu and Murphy's (1993) mainland China study. In fact, the opposite was observed. A third international study of self-ratings was conducted in a Hong Kong-based airline. In this study, Furnham and Stringfield (1994) found that the self-ratings and supervisor ratings did not differ for the Chinese (i.e., Cantonese) sample. In sum, three different studies of Chinese self-raters have yielded remarkably different results.

It appears then that further study of possible differences in self-ratings between national cultures is warranted. The question of whether a collectivist cultural orientation produces a modesty or leniency effect or, even no effect at all on self-ratings, would benefit from a study done in other collectivist countries than those using Chinese subjects. Two collectivist countries, Poland and Russia, are addressed in the present study.

Research in Poland and Russia

Based on Kolman, Hofstede, Noorderhaven, and Dienes' (1999) work, Hofstede (2001) reported that Poland's individualism-collectivism score is 60. In the original Hofstede IBM study reported in Hofstede (2001), the U.S. was classified as highly individualistic because of its index score of 91. The comparison of the two country's scores suggests that Poland is more collectivist than is the U.S. Based on Hofstede, Kolman, Nicolescu, and Pajumaa's (1996) published work as well as raw data from unpublished studies reported in Hofstede (2001), Hofstede concluded that Russia's collectivism-individualism score is 39 which is far below the U.S. and even Poland, thereby placing Russia in the highly collectivist category.

A thorough search of both the Polish and Russian literatures yielded two Polish studies that relate to the self-rating process, while no self-rating studies were discovered in the Russian literature. Both Polish studies will now be discussed. Dabul, Wosinka, Cialdini, Mandel, and Dion (1997) found that the university-age Polish subjects assigned higher workplace rewards to the Polish employee making a highly modest self-presentation versus a more boastful self-presentation. Thus, we see a strong preference among this

sample of university-age Poles for self-effacement when presenting oneself to others. Dabul et al. (1997) attributed this finding to the strong collectivist cultural values present in Poland.

Schultz (2001) reported the tendency for some Polish self-raters employed by a chemical company to rate their own performance lower than their bosses. That is, 120 of the 600 managers in her sample rated their own performance consistently lower across all dimensions than did their supervisor. Conversely, the rest of the managers in her study either rated their own performance at the same level as their supervisor (i.e., 294 managers fell in this category), higher than their supervisor (i.e., 54 managers were in this grouping), or higher than the supervisor on some dimensions but lower on other dimensions (i.e., 132 in this category). Even though the data provided by Schultz (2001) suggest that 20% of the Polish self-ratings in her study reflect some form of harshness bias while another 9% reflect a leniency bias, no firm conclusions can be drawn because a statistical test for mean differences was not conducted.

Dabul et al.'s (1997) self-presentation findings and the preliminary results of Schultz (2001) suggest that Polish self-ratings may support Farh et al.'s (1991) proposition that collectivism is related to modesty bias in self-ratings. No similar studies were found for Russian employees, but in line with Farh et al.'s proposition, collectivist Russia would be expected to demonstrate similar results to those found in Poland. Accordingly, the following hypothesis is proposed:

Hypothesis 1: Employee self-ratings of workplace behavior in collectivist cultures, such as Poland and Russia, will be lower than supervisory behavioral ratings of that employee.

Research in the United States

Contrary to the Taiwanese experience, appraisal research conducted in the U.S. has shown that subordinates tend to evaluate their own performance at a higher level than their supervisors (Harris & Schaubroeck, 1988). This leniency finding in U.S. self-rating research is fairly strong in that it has been found for various rating instruments and for various rating purposes.

Farh et al. (1991) argued that an individualistic cultural orientation (e.g., Hofstede, 2001) provides a cultural context for the inflated self-ratings that are observed in the individualistic U.S. Individual accomplishment and self-sufficiency characterize the individualistic dimension. According to Farh et al. (1991), such individualistic values are likely contributors to the leniency effect found in U.S. self-ratings. Providing further support for this, Markus and Kitayama (1991) concluded that persons in independent cultures tend to be focused more on the self instead of others, thereby yielding a more self-enhancing posture. Thus, the individualistic orientation of U.S. workers provides an underlying explanation for leniency in U.S. self-ratings.

Based on prior U.S. results and the individualism reasoning, it is likely that lenient self-ratings will surface in our U.S. sample. However, the self-rating studies included in

Harris and Schaubroeck's (1988) meta-analysis were conducted in the pre-downsizing era in the U.S. Might harsher self-ratings occur because U.S. workers do not feel overly productive after experiencing the turbulent downsizings and staff shortages in today's typical U.S. firm? Nonetheless, this notion of harsher self-ratings in U.S. firms is only speculative; it is more likely, considering the literature, that U.S. self-ratings will reflect leniency. Accordingly, the following hypothesis is put forth:

Hypothesis 2: Employee self-ratings of workplace behavior in an individualistic culture, such as the United States, will be higher than the supervisory ratings of that employee.

Method

Participants

Ninety-one U.S. participants were employed in their respective organizations while enrolled in adult evening programs at a local university in Western New York. Another seven employees worked full-time in a small U.S. technology firm. The supervisor of each employee also participated in this study. The Polish sample is made up of 99 employee-supervisor dyads from different firms located in Warsaw. These Polish employees were enrolled in an adult program in a major Polish university while working in their respective organizations. The Russian sample is made up of 86 dyads from different firms located in and around the city of Krasnoyarsk in central Russia. The Russian participants were recruited from among current adult students and former students of a local university, who were employed in their respective firms. Although most of the participants in this study were recruited through their connection with a local university, they should not be considered inexperienced employees. The description of the U.S., Polish, and Russian samples in Table 1 indicates that the length of tenure of the U.S., Polish, and Russian participants in their firms is 7.49, 3.91, and 9.29 years, respectively. Other demographic statistics for each country's sample are provided in Table 1.

Measures

The questionnaire administered to employees and supervisors included various demographic items (i.e., length of service in the job and firm, gender, and managerial status) and five items assessing the target employee's behavior in various areas. Employees were asked to evaluate their own behavior in five areas: creativity, motivation, risk taking, initiative, and assertiveness. The same items were included, with appropriate minor changes in wording, in the supervisor's questionnaire. The exact wording of these behavioral items are: "My (his/her) level of creativity in the job is generally very high," "I (s/he) display(s) a high level of motivation and energy in my (the) job," "I (s/he) am not afraid of taking risks and trying new things in my (the) job," "I (s/he) take immediate action to resolve work-related problems as they emerge," and "I (s/he) volunteer my (his/her) work-related views without waiting to be asked." A five-point Likert scale (1 = "strongly agree," ..., 5 = "strongly disagree") was used as anchors in the previously mentioned behavioral scales. To avoid confusion when interpreting data,

all ratings were reverse-scored. The Polish and Russian questionnaire items and rating instructions were translated into their respective language. In line with Pavett and Morris (1995), all measures and instructions were then back-translated into English by a different interpreter in order to insure translation accuracy.

Procedures

Employees in this study completed their self-ratings and then returned their responses in person or by mail. They were also asked to deliver a similar questionnaire to their immediate supervisor. In the instructions, the supervisor was told that they could return their completed ratings by mail or have their employee return their sealed responses. The estimated response rate for the U.S. sample is 35% and 50% for the Polish and Russian samples. All participants were instructed that their ratings were to be used for research purposes, which is consistent with the instructions given in other international self-rating studies, such as Farh et al. (1991), Yu and Murphy (1993), and Furnham and Stringfield (1994) as well as U.S. studies reported in Harris and Schaubroeck's (1988) meta-analysis.

Results

The two hypotheses in this study posit that the Polish and Russian self-ratings will be lower than the supervisor ratings assigned to these same employees whereas the U.S. self-ratings will be higher than the supervisor ratings. Similar to the statistical procedures used in Farh et al. (1991) and Yu and Murphy (1993), we conducted a 3 x 2 x 5 (U.S./ Poland/ Russia X self-ratings/supervisor ratings X creativity/motivation/risk taking/initiative/assertiveness) analysis of covariance (ANCOVA) to test whether self-ratings differed from the supervisor ratings by country and perhaps by behavioral dimension. Rating source (self-ratings vs. supervisor ratings) and dimension (creativity vs. motivation vs. risk taking vs. initiative vs. assertiveness) are within-subjects variables, while country (U.S. vs. Poland vs. Russia) was treated as a between-subjects variable.

Table 1. Descriptive Statistics for the Sample in Each Country

	<u>U.S.</u>	<u>Russia</u>	<u>Poland</u>
Focal Employees:			
Female	58%	62%	70%
Management Position	37%	34%	18%
Average tenure in current job	3.75 yrs.(SD = 3.32)	5.38 yrs.(SD = 5.66)	2.53 yrs.(SD = 3.26)
Average tenure in the firm	7.49 yrs.(SD = 6.26)	9.29 yrs.(SD = 9.07)	3.91 yrs.(SD = 4.70)
Supervisors:			
Female	38%	47%	53%
Average tenure in current job	4.33 yrs.(SD = 3.32)	5.58 yrs.(SD = 5.54)	4.61 yrs.(SD = 5.00)
Average number of years supervising focal employee	3.07 yrs.(SD = 2.94)	6.86 yrs.(SD = 6.94)	2.31 yrs.(SD = 2.91)

The ratings assigned to the target employee on the five dimensions served as dependent measures in this study. To control for extraneous effects of sample differences on the self-ratings, three demographic variables (i.e., employee gender, managerial status, and job tenure) served as covariates. As shown in Table 2, the ANCOVA results indicated that the three-way interaction (country X rating source X dimension) was not significant ($p = .20$), suggesting that ratings from the two rater sources did not differ by country and by behavioral dimension. The results in Table 2 did show, however, a significant two-way country X rating-source interaction, $p < .01$. As expected, the pattern of the ratings made by each country's raters differed markedly (see Table 3). To determine the exact nature of the significant two-way interaction, paired t-tests, contrasting the two rating sources in each country, were conducted.

Table 2. Rating Source x Country x Behavioral Dimension ANCOVA

Source of Variation	df	MS	F	eta ²
Rating source	1	4.87	3.60	.01
Rating source x tenure	1	.20	.15	.00
Rating source x managerial status	1	2.18	1.61	.01
Rating source x gender	1	1.81	1.34	.01
Rating source x country	2	8.26	6.10**	.05
Error (Rating source)	253	1.35		
Dimension	4	.22	.50	.00
Dimension x tenure	4	.65	1.48	.01
Dimension x managerial status	4	.24	.55	.00
Dimension x gender	4	.64	1.46	.01
Dimension x country	8	1.32	3.02**	.02
Error (Dimension)	1012	.44		
Rating source x dimension	4	.04	.12	.00
Rating source x dimension x tenure	4	.24	.70	.00
Rating source x dimension x managerial status	4	.07	.22	.00
Rating source x dimension x gender	4	.14	.41	.00
Rating source x dimension x country	8	.47	1.39	.01
Error (Rating source x Dimension)	1012	.34		
Employee's job tenure	1	1.37	.67	.00
Managerial status	1	16.33	7.99**	.03
Employee's gender	1	4.00	1.96	.01
Country	2	44.37	21.70**	.15
Error	262	2.76		

* $p < .05$. ** $p < .01$.

As can be seen in Table 3, the direction of mean differences and the corresponding paired t-test results indicated that the Polish self-ratings are significantly lower than the Polish supervisor ratings on the motivation ($t_{98} = -3.71, p < .001$), risk taking ($t_{98} = -2.80, p < .01$), and assertiveness dimensions ($t_{98} = -3.55, p < .01$). These findings provide some support for Hypothesis 1. The mean difference between the Polish self-rating and the supervisory rating on the creativity dimension barely missed reaching the conventional significance level, ($t_{98} = -1.93, p = .06$). Note also that the direction of the difference in creativity means is consistent with this study's first hypothesis. The results show that no difference was detected for the initiative dimension, ($p = .62$). When each Polish rater's five behavioral ratings were combined into an overall composite, the direction of the composite mean difference and the corresponding paired t-test result (see Table 3) revealed that the average Polish self-rating across the five behavioral dimensions is significantly lower than the average Polish supervisor rating, ($t_{98} = -3.67, p < .001$). Taken as a whole, these Polish results support Hypothesis 1.

Table 3. Mean Self- and Supervisory Ratings

<u>Country/Dimension</u>	<u>Self-Rating</u>		<u>Supervisory Rating</u>	
	<u>M</u>	<u>SD</u>	<u>M</u>	<u>SD</u>
United States (N = 98):				
Creativity	3.76**	.83	4.02	.72
Motivation	4.18*	.68	4.39*	.76
Risk Taking	4.13	.83	4.07	.79
Initiative	4.24	.58	4.29	.66
Assertiveness	4.08	.71	4.16	.83
Average U. S. Rating	4.08	.49	4.19	.55
Russia (N = 86):				
Creativity	3.50	.75	3.50	.86
Motivation	3.89	.62	3.71	.95
Risk Taking	3.84**	.73	3.52**	.97
Initiative	4.06*	.64	3.84*	.92
Assertiveness	3.72	.72	3.68	1.01
Average Russia Rating	3.80	.49	3.65	.74
Poland (N = 99):				
Creativity	3.63	.88	3.85	.75
Motivation	3.62**	.91	4.01	.75
Risk Taking	3.47**	.89	3.81	.94
Initiative	3.89	.70	3.94	.83
Assertiveness	3.51	1.03	3.89**	.90
Average Poland Rating	3.63	.60	3.90**	.62

* $p < .05$ level. ** $p < .01$ level.

Russian ratings were included in this study to assess whether the expected harshness effects of the Polish self-ratings would generalize to another collectivist Eastern European country. Thus, the Russian ratings provided a necessary supplement to the Polish self-ratings in terms of determining whether the cause of the self-rating harshness effect is rooted in the "collectivist cultural explanation" or whether it is unique to Poland. With the exception of the risk-taking and initiative ratings, all other comparisons of Russian self-ratings with Russian supervisory ratings were not significant (Table 3). On two of the five dimensions, the Russian self-raters evaluated their own performance higher than did their supervisors, ($t_{85} = 2.89$, $p < .01$ on risk taking and $t_{85} = 2.08$, $p < .05$ on initiative). When each Russian rater's five ratings were combined into an overall composite, the paired t-test result (Table 3) indicated that the average Russian self-rating across the five dimensions was not significantly different from the average Russian supervisory rating, ($t_{85} = 1.85$, $p = .07$), although it just missed the conventional level of significance. These Russian results, showing lenient self-ratings on two of the five behavioral dimensions, suggest that the harshness effect detected in the Polish self-ratings did not generalize to the Russian self-ratings. Hypothesis 1 is not supported by these results.

To further interpret the two-way country X rating-source interaction, paired t-tests contrasting the two sets of U.S. ratings were also conducted. Contrary to this study's second hypothesis, all but two of the U.S. self-rating/supervisor rating comparisons were not significant. Specifically, the results in Table 3 indicate that the U.S. self-rating was equivalent to the supervisory rating on risk taking ($t_{97} = .59$, $p = .56$), on initiative ($t_{97} = -.50$, $p = .62$), and on assertiveness ($t_{97} = -.76$, $p = .45$). On the other hand, the U.S. self-ratings on both the creativity and motivation dimensions were significantly lower than the supervisory ratings ($t_{97} = -2.68$, $p < .01$ on creativity and $t_{97} = -2.17$, $p < .05$ on motivation), suggesting that these self-ratings reflect a harshness bias instead of the hypothesized leniency effect. When each U.S. rater's five ratings were combined into a composite, the t-test result (Table 3) revealed that the average U.S. self-rating and the average U.S. supervisory rating across the five behaviors are not significantly different, ($t_{97} = -1.59$, $p = .12$). Taken as a whole, these U.S. results do not support this study's second hypothesis.

Discussion

Our research findings dispel support for Farh et al.'s (1991) modesty-collectivism proposal. The self-ratings made in two highly collectivist countries, Russia and Poland, should have yielded more consistent results, if the modesty notion were valid. Considering Yu and Murphy's (1993), Furnham and Stringfield's (1994), and this study's results, the consensus seems to be that Farh et al.'s (1991) modesty-collectivist effect on self-ratings is without merit. If collectivism cannot explain the harshness effect detected in Polish self-ratings, then what is the underlying explanation for this finding? An attempt should be made to better understand the nature of this severity effect. One explanation could be based on Marody's (1988) self-deprecation suggestion. The Polish

commoner, in her view, places a "bonus on mediocrity," making sure that one's self-presentation fits the image of the "average" Pole instead of the elite Pole.

If Marody (1988) is correct about the self-deprecating Pole, then what accounts for this tendency? Plausible explanations might be rooted in economic, historical, or even religious realities. Similar to Taiwan, the severe self-ratings found in the Polish sample may be reflective of a relatively small country. The harsh economic conditions in a beleaguered Poland, that has historically been Europe's political pawn, may abet the deflated self view of workplace performance. In line with Adam's (1963) equity theory, economically under-rewarded and historically underappreciated Polish workers may return to a state of equity by psychologically devaluing their worth on the job. Lastly, Poland's strong religiosity stemming from its Catholic faith could be playing a role in the self-rater's propensity to rate with humility.

Even though this harshness effect does not appear to generalize to Russian self-ratings, it should be determined whether it extends to other neighboring Central-Eastern European countries. A recent Economist article (Lucas, 1997) referred to this self-deprecating phenomenon in Central-Eastern Europe. One researcher (Feher, 1991) reported that discussing one's individual performance in the workplace is considered impolite in this region of Europe. Uncovering the extent to which Russian self-ratings differ from other Central-Eastern European countries and understanding the underlying reason for such differences could be pursued.

Our U.S. results (i.e., equivalent ratings on three of five dimensions and a severity effect on the creativity and motivation dimensions) may, at first, seem inconsistent with the typical U.S. self-rating leniency finding. At the same time, just as the validity of Farh et al.'s modesty-collectivist proposition has been seriously questioned by our research, perhaps this leniency effect identified for countries with individualistic values has more situational factors than previously thought. Could our unanticipated U.S. self-rating results on the creativity and motivation dimensions be reflecting a marked shift in the U.S. work environment? Many of the past self-rating studies in the U.S. were conducted in the pre-downsizing time period whereas the present study was conducted in the rapid-change, downsizing era. Might the harsh self-ratings found in this study be reflecting a weary U.S. worker that does not feel overly creative or motivated after experiencing extremely stressful working conditions (e.g., staff shortages) that come with the downsizing?

We cannot rule out the possibility that our sampling approach accounts for our unexpected findings. Because the participants were directly or indirectly recruited from a university setting, our samples may have an unrepresentative number of younger, educated employees, but since the average organizational tenure of our U.S., Russian, and Polish participants is 7.49, 9.29, and 3.91 years, respectively, these self-raters should not be viewed as inexperienced. Nonetheless, our samples may not be representative of each country's workforce.

References

- Adams, J.S. (1963). Toward an Understanding of Inequality. *Journal of Abnormal and Social Psychology*, 67, 422-436.
- Dabul, A.J., Wosinska, W., Cialdini, R.B., Mandal, E., & Dion, R. (1997). Self-Presentational Modesty Across Cultures: The Effects of Gender and Social Context in the Workplace. *Polish Psychological Bulletin*, 28, 295-306.
- Elenkov, D.S. (1998). Can American Management Concepts Work in Russia?. *California Management Review*, 40, 133-156.
- Farh, J.L., Dobbins, G.H., & Cheng, B.S. (1991). Cultural Relativity in Action: A Comparison of Self-Ratings made by Chinese and U.S. Workers. *Personnel Psychology*, 44, 129-147.
- Feher, J. (1991). Kritikus Pontok a Teljesitmenyertekelesi Rendszerek Hatekony Alkalmazasaban. *Humanpolitikai Szemle*, 78, 13-22.
- Furnham, A., & Stringfield, P. (1994). Congruence of Self and Subordinate Ratings of Managerial Practices as a Correlate of Supervisor Evaluation. *Journal of Occupational and Organizational Psychology*, 67, 57-67.
- Harris, M.M., & Schaubroeck, J. (1988). A meta-analysis of self-supervisor, self-peer, and peer-supervisor ratings. *Personnel Psychology*, 41, 43-62.
- Hofstede, G., Kolman, L., Nicolescu, O., & Pajumaa, I. (1996). Characteristics of the Ideal Job Among Students in Eight Countries. In H. Grad, A. Blanco, and J. Georgas (Eds). *Key Issues in Cross-Cultural Psychology* (pp. 199-216). Lisse, Netherlands: Van Gorcum.
- Hofstede, G. (2001). *Culture's Consequences: Comparing Values, Behaviors, Institutions, and Organization Across Nations*. Thousand Oaks, CA: Sage Publications.
- Kolman, L., Hofstede, G., Noorderhaven, N., & Dienes, E. (1999). Work-Related Values and Cooperation in Europe. In *Konferencje PEF CZU, Agrarni Perspektivy VIII [Proceedings of the Conference on Agrarian Perspectives]* (pp. 884-890). Prague: PEF/CZU.
- Lucas, E. (1997). Doing Business. *Economist*, 345 (Nov. 22, 8044), S13-S15.
- Markus, H.R., & Kitayama, S. (1991). Culture and the Self: Implications for Cognition, Emotion, and Motivation. *Psychological Review*, 98, 224-253.
- Marody, M. (1988). Antinomies of Collective Subconsciousness. *Social Research*, 55, 97-110.
- Schultz, A. (2001). Ocena i Samoocena Kadry Kierowniczej. *Humanizacja Pracy*, 199, 59-66.
- Yu, J., & Murphy, K.R. (1993). Modesty Bias in Self-Ratings of Performance: A Test of the Cultural Relativity Hypothesis. *Personnel Psychology*, 46, 357-363.

Contact email addresses:

costigan@sjfc.edu insingrc@oneonta.edu gkranas@sci.psych.uw.edu.pl
 ilter@sjfc.edu berman@sjfc.edu kureshov@stu.krasnoyarsk.su