Self-Other Rating Agreement and Humility, Narcissism, and Academic Performance

Cade Bleashka and Darrin Kass, Ph.D.

Department of Management and International Business
Bloomsburg University of Pennsylvania

ABSTRACT

The purpose of this paper was to examine the relationship between self-other rating agreement (SOA) and humility, narcissism, and academic performance. Three SOA categories were created from a sample of 226 undergraduate students (111 women and 115 men): 1) Self-aware, individuals whose self-ratings matched observer ratings, 2) Underraters, those whose self-ratings were lower than observer ratings, and 3) Overraters, individuals whose self-ratings were higher than other ratings. ANOVA revealed that GPA varied by SOA, with underraters showing the highest GPA (3.11), followed by accurate (3.05), and then overraters (2.86). One of the intriguing results is that underraters displayed both the lowest narcissism and humility scores. This research sheds light on the individual differences associated with SOA and reinforces its impact on performance in various contexts.

Keywords: rating accuracy; self-awareness; performance

Introduction

Research has consistently found that self-other rating agreement (SOA) impacts organizational outcomes. Those individuals who rate themselves most congruently with their others (team members, co-workers, supervisors) demonstrate the most positive organizational outcomes (Yammarino and Atwater 1997), as compared to individuals who either over or underestimate themselves. The outcomes of SOA have been well-established in the existing body of literature (e.g., Atwater et al. 1998; Atwater and Yammarino 1992; McKee et al. 2018; Owens et al. 2015; Rego et al. 2020; Ridge and Ingram 2017; Swain 2018), but the relationship between SOA and academic performance has not been investigated. Further, less is understood about SOA as a concept. While it is represented by the difference between self and other ratings, little is understood about what this difference actually represents. Fleenor et al. (2010) noted that it may be a constellation of traits, but research to date has not examined what these traits may be. This research sought to fill this void by examining the relationship between self-other (dis)agreement and academic performance and two traits likely associated with underrating and overrating: humility and narcissism, respectively.

Yammarino and Atwater (1993) define SOA as the relationship between how individuals rate their own performance and how others rate that individual's performance. If the individual and the other rate the individual in a similar fashion, this is considered an accurate rating. Alternatively, if there is disagreement between how the individual and other rate, two additional situations occur: underrating and overrating, individuals who rate themselves lower or higher, respectively, than their others.

These SOA categories have been linked to numerous relevant organizational outcomes, including performance (e.g., Rego et al. 2020; Yammarino and Atwater 1993), leadership (e.g., Atwater and Yammarino 1992; McKee et al. 2018), and

even ethical behavior (e.g., Kass et al. 2021; Kuenzi et al. 2019). The most positive outcomes have been associated with accurate raters, and inaccurate ratings tend to have negative outcomes, whether or not they are associated with underrating or overrating (Yammarino and Atwater 1993).

SOA and performance outcomes. Studies have consistently shown that performance varies by SOA, with higher levels of performance linked to accurate raters. Yammarino and Atwater (1993) found that accurate raters are "more successful (regardless of how 'success' is defined) than either over-estimators or under-estimators" (p. 240). Accurate raters have better workplace attitudes, show higher commitment to their work, make more effective decisions, achieve more promotions, and have the highest levels of performance and success in their positions (Yammarino and Atwater 1997). Underrating has typically been associated with mixed results in the literature, making their performance outcomes somewhat difficult to predict (Yammarino and Atwater 1997). However, the majority of the results tend to be positive, particularly when performance ratings are from their subordinates (Atwater et al. 1998; Sosik and Megerian 1999). Underraters are perceived as better leaders, have more engaged employees, are more likely to possess leadership strengths (Zenger and Folkman 2015), and develop the highest levels of trust and organizational commitment amongst their subordinates (Sosik 2001). Lastly, overrating is predominantly associated with poor performance outcomes. Overraters tend to have more negative attitudes, lack awareness of their strengths and weaknesses, and are less likely to utilize training and development opportunities (Woo et al. 2008). Atwater et al. (1998) noted that overraters tend to ignore developmental feedback and, so, fail to address performance deficiencies. Based on prior studies, we expect that academic performance will vary by SOA. Accurate raters should have the highest GPA followed by underraters and then overraters.

H1: Accurate raters will exhibit the highest GPA scores, followed by underraters and overraters.

SOA and individual differences. Research has shown that a variety of demographic factors and personality traits have been associated with the SOA categories. Fleenor et al. (2010) noted that male, older, less educated, and non-white individuals are more likely to overrate. McKee et al. (2018) found that accurate raters were higher in agreeableness and neuroticism, and that overrating was associated with higher levels of conscientiousness. Predictions associated with personality and SOA tend to be complex because "observable traits will be more strongly related to others' ratings, and less observable traits (more internally manifested) will be more strongly related to self-ratings" (p. 291).

We would also expect SOA to be theoretically linked to both humility and narcissism. These traits are observable, and therefore are likely to be related to both self and other ratings. Underraters provide lower self-scores than those provided by others, which might suggest the behavior of humble individuals. Humble individuals see themselves as a work in progress (Owens and Hekman 2012) and feel as though they fall short of a standard they cannot reach (Ou et al. 2014). They believe they need to develop, evolve, and improve, and this drive may result in higher performance. Ridge and Ingram (2017) note that "modesty positively impacts career success and upward mobility" (p. 1283) and Rego et al. (2020) suggests that humility is a distinguishing feature of many well-respected leaders.

Overrating may be a form of narcissism. Self-enhancement is a central aspect of narcissism (Grijalva and Zhang 2016), and narcissists tend to overrate their

attributes and abilities, such as leadership effectiveness, intelligence, physical attractiveness, openness, and honesty. Gebauer et al. (2012) noted that narcissists see themselves as smarter than non-narcissists, although objective data does not necessarily support that they are actually smarter. Owens et al. (2015) suggests that organizations with high numbers of narcissistic leaders show more lapses in ethical conduct and issues tend to arise more often and to a greater degree. On the other hand, humility is shown to have a more positive effect for organizations than narcissism has proven. Swain (2018) finds that underraters (those who are humbler) perform the best in virtual groups.

Therefore, we expect that humility and narcissism scores will vary by SOA. We anticipate that underraters will have the highest scores on humility and that overraters will have the highest narcissism scores.

H2: Underraters will have significantly higher scores on humility than accurate and overraters.

H3. Overraters will have significantly higher levels of narcissism than accurate and underraters.

METHODS

Participants

We collected data from students in the College of Business enrolled at a regional university in the Pennsylvania State System of Higher Education. Data collection was approved by the BU Institutional Review Board (IRB) under study IRB# 2017-31. Courses were selected that established semester-long teams that worked together on graded assignments over the course of the semester. Only teams whose membership remained intact over the course of the semester were included in the study. The final sample consisted of 44 teams with 226 participants: 111 women and 115 men. Ages ranged from 20 to 47

years, with an average age of 22 years. The majority of participants were between the ages of 20 and 23 (92.1%).

Measures

Rating accuracy. At the time of data collection, participants were asked to rate themselves and their teammates on five dimensions of effective team behavior on a Likert scale from 1 "none at all" to 5 "a great deal." These dimensions included: initiative, communication, cooperation, preparedness, and contribution. Self-ratings (SR) were established by the mean of self scores on the five dimensions. Other ratings (OR) included the mean scores of all teammates on the same five dimensions. The classification for over, under, or self-aware rating was made using a procedure well established in the literature by Atwater and Yammarino (1992). Rating accuracy was defined as the difference score between the self-ratings of participants and the scores provided by their teammates. SOA was established by comparing the average of teammate ratings (OR) with the average of self ratings (SR). Participants whose difference scores were half a standard deviation above the mean were considered overraters. Participants whose difference score was half a standard deviation below the mean were labeled underraters. Lastly, individuals with half a standard deviation above or below the mean were categorized as accurate. In our sample of 226, 74 were classified as underraters, 86 as accurate raters, and 66 as overraters. The mean scores of both self and other ratings across the SOA categories are presented in Table 1.

Grade Point Average (GPA). GPA was included as a measure of the academic performance of participants at the time of data collection. This was assessed from official student records. GPA was measured on a 4.0 scale, and the range for GPA included in this sample was restricted to a range of 2.0 - 4.0 with an overall mean of 3.01.

Table 1. *Descriptive Statistics and Correlations of Variables Under Study*

Variable	Mean (SD) Under	Mean (SD) Accurate	Mean (SD) Over	1	2	3	4
1. SR	3.72 (.39)	4.40 (.42)	4.50 (.49)	__			
2. OR	4.73 (.24)	4.69 (.34)	4.03 (.70)	.25**	__		
3. Humility	4.10 (.46)	4.37 (.41)	4.52 (.48)	.50**	.00	__	
4. Narcissism	13.49(6.58)	17.65(8.28)	17.85(7.35)	.36**	.07	.15*	__
5. GPA	3.11(.49)	3.05(.40)	2.86(.48)	.03	.33**	-.08	-.09

Note. SR = Self-Rating; OR = Other-Ratings
 * $p<.05$, ** $p<.01$

Humility. Humility was measured with the nine-item expressed humility scale (Owens et al. 2013) on a Likert scale ranging from 'strongly disagree' to 'strongly agree.' Expressed humility is a strengths-based measure of humility and includes the following dimensions: 1) a willingness to view oneself accurately, 2) an appreciation of the strengths and contributions of others, and 3) teachability.

Narcissism. Narcissism was assessed with the Narcissistic Personality Inventory (NPI), which is the most widely used measure of narcissism (Raskin and Hall 1979). The scale uses 40 forced-choice items where participants choose responses that are either narcissistic (e.g., 'I can live my life anyway I want to') or non-narcissistic (e.g., 'People can't always live their lives in terms of what they want'). The 40 items are summed together based on narcissistic choices, and higher scores indicate higher levels of narcissism.

RESULTS

Descriptive statistics and correlations are displayed in Table 1. All dependent variables had skewness and kurtosis scores between –1 and 1 and, therefore are considered normal. There were no significant differences on the dependent variables based on age or gender save for narcissism, where men (M = 18.68) reported being significantly more narcissistic than women (M = 14.03). Self-ratings were positively correlated with humility and narcissism, indicating that higher humility and narcissism scores were associated with higher self-ratings. Interestingly, other-ratings were positively correlated with GPA, but self-ratings were not. It was somewhat surprising to see a positive relationship between humility and narcissism, although this relationship was weak (0.15).

ANOVA was used to test for differences in participant's' GPA, humility, and narcissism between the three SOA groups (Table 2). For all three variables, SOA had a significant effect ($p<.01$). Overall, SOA is a significant predictor of GPA, narcissism, and humility. These results are both statistically and practically significant since these effect sizes are considered medium (GPA and Narcissism) and large (Humility).

Table 2: *Differences by group in GPA, Humility, and Narcissism*

Variable	F (2,225)	η2	Under (U)		Over (O)		Acc (A)	
			M	SD	M	SD	M	SD
GPA	5.38*	.05	3.11[b]	.49	2.86[bc]	.48	3.05[c]	.40
Humility	16.38*	.13	4.10[ab]	.46	4.52[b]	.48	4.37[a]	.41
Narcissism	8.03*	.07	13.49[ab]	6.58	17.85[b]	7.35	17.64[a]	8.28

Note: [a] =significant difference between Under and accurate. [b] = significant diff between under and over, [c] = significant difference between accurate and over
*p<.01

In order to examine differences among the rating groups and test our hypotheses, we used the Tukey HSD procedure (Table 3). For GPA, the underraters (3.10) were significantly higher than overraters (2.86). Further, overraters had significantly lower GPAs than accurate raters. Underraters did not achieve a significantly higher GPA than the accurate (3.05). Thus, hypothesis 1 is not supported. For both narcissism and humility, a similar pattern of results is found. Underraters significantly differ from accurate and overraters, but there are no differences between accurate and overraters. For narcissism, overraters were the most narcissistic (17.85), followed by accurate (17.65) and underraters (13.4). This supports hypothesis 2. Somewhat surprisingly, this pattern follows for humility, with underraters having the lowest scores on humility (4.10) as compared to accurate (4.37) and overraters (4.52). Therefore, Hypothesis 3 is not supported. For both humility and narcissism, all differences were significant except for those between accurate and overraters, although it did approach significance for humility (p = 0.08).

Table 3: *Pairwise Comparison by Tukey Method*

Variable	Pair	Mean Difference	95% LB	95% UB
GPA	Under – Acc	.06	-.12	.23
	Under – Over	.24*	.06	.43
	Acc – Over	.19*	.01	.36
Humility	Under – Acc	-.27*	-.44	-.10
	Under – Over	-.42*	-.60	-.25
	Acc– Over	-.16	-.33	.02
Narcissism	Under – Acc	-4.16*	-6.97	-1.36
	Under – Over	-4.36*	-7.35	-1.37
	Acc– Over	-.20	-3.09	2.69

Note. *: *p<.05*

DISCUSSION

The purpose of this paper was to determine if differences between self and other ratings were related to academic performance and personality characteristics amongst college students. SOA has been linked to numerous important outcomes, but no research to date has explored its relationship to GPA. Further, little is known about SOA as a construct, so we sought to explore two personality characteristics—humility and narcissism—that could likely affect SOA. Overall, our results showed that there were significant differences among under, over, and accurate raters on GPA, humility, and narcissism. These findings suggest that individual differences in self and other ratings are personally and psychologically meaningful, and worthy of study.

In our study, we found that academic performance varied by SOA—underraters and accurate raters had significantly higher GPAs than overraters. It is worth noting that underraters had higher GPAs than accurate raters (3.10 vs 3.05), but this difference did not reach significance. This is relatively consistent with previous literature, particularly with overraters showing the lowest levels of performance across the board.

As expected, overraters exhibited the highest levels of narcissism. The higher than warranted self-scores may reflect the overconfidence associated with narcissists (Campbell et al. 2004). This is an important finding because narcissism is a broad and stable personality trait, indicating that SOA may reflect a stable, general tendency to enhance, diminish, or accurately rate oneself in comparison to others. It also sheds light on the process of overrating in that those who overrate may see themselves in a better light than is warranted and desire to both enhance and maintain these inflated estimates (Campbell et al. 2004). Both narcissism and overrating have been associated with problematic outcomes, particularly with performance. Brenkert (2019) discovered that non-humble (narcissistic) leaders create less ethically-minded work environments due to employees not wanting to speak up—resulting in a deficiency in performance. In an academic setting, Gebauer et al. (2012) found that the actual knowledge base of narcissists and non-narcissists differs insignificantly, but narcissists present themselves as having greater "intelligence, course grades, creativity, and academic knowledge" (p. 854). Because of this perception, narcissists do not seek to improve, leading to lower performance levels.

The most surprising result of our study is that overraters, not underraters, had the highest levels of humility. Although unexpected, if SOA is a stable characteristic, it follows that underraters and overraters are consistent in their ratings across self-reports of personality. Although underraters may be humble, they see themselves as a constant work in progress. Therefore, they may believe they could be humbler and rate themselves accordingly. Likewise, overraters may overrate themselves on all of their characteristics, and therefore end up with the highest scores on humility despite not actually being humble. If this is the case, it logically follows that overraters would provide high self-scores on humility—because they think of themselves higher in all instances.

The results of this study are meaningful because they provide a better understanding of SOA. First, we found that SOA predicts objective, academic performance. This adds to the large body of literature of the relationship between SOA and meaningful performance outcomes. Interestingly, underraters had the highest GPAs, followed by accurate and then overraters. The largest performance differences were between underraters and overraters. Second, our results clearly indicate that self and other ratings should

be used in research on individual differences. Both self and other ratings had different relationships with the dependent variables, and the use of self-ratings alone would lead to very different interpretations. In our study, the most narcissistic group (overraters) also rated themselves as the humblest. Using only self-scores in this instance would lead to very flawed conclusions.

Future Directions and Limitations

In consideration of the reliability of the results, this study did have several limitations. We included only undergraduate students in semester-long teams that met for three hours per week during the semester. Although these teams worked collaboratively on projects, it was in an academic sense, so it is unknown if our results of the study would generalize to an organizational environment. Additionally, the sample is from a fairly small, public, and regional university where many of the students are Caucasian and local to the area. This study would need a larger, more diverse sample to be confident in the generalizability of the results. Fleenor et al. (2010) noted that SOA scores are affected by culture, age, and gender.

While Yammarino and Atwater (1993) argued that rating accuracy leads to the most positive outcomes, our study found underraters had the highest GPAs. Further study would be necessary to determine if underrating always equals high performance, or if it is just so in certain circumstances. Lastly, more research is needed to determine if SOA is actually a stable personality characteristic. We found significant linkages between SOA and humility and narcissism, but additional work is needed to determine if these differences are stable across time and additional individual differences.

LITERATURE CITED

Atwater, L.E., and F.J. Yammarino. 1992. Does self-other agreement on leadership perceptions moderate the validity of leadership and performance predictions? *Personnel Psychology* 45(1): 141-164. doi: 10.1111/j.1744-6570.1992.tb00848.x

Atwater, L.E., C. Ostroff, F.J. Yammarino, and J.W. Fleenor. 1998. Self-other agreement: Does it really matter? *Personnel Psychology* 51(3): 577-598. doi: 10.1111/j.1744-6570.1998.tb00252.x

Brenkert, G.G. 2019. Mind the gap! The challenges and limits of (global) business ethics. *Journal of Business Ethics* 155(4): 917-930. doi: 10.1007/s10551-018-3902-6

Campbell, W.K., A.S. Goodie, and J.D. Foster. 2004. Narcissism, confidence, and risk attitude. *Journal of Behavioral Decision Making* 17(4): 297-311. doi: 10.1002/bdm.475

Fleenor, J.W., J.W. Smither, L.E. Atwater, P.W. Braddy, and R.E. Sturm. 2010. Self-other rating agreement in leadership: A review. *The Leadership Quarterly* 21(6): 1005-1034. doi: 10.1016/j.leaqua.2010.10.006

Gebauer, J.E., C. Sedikides, B. Verplanken, and G.R. Maio. 2012. Communal narcissism. *Journal of Personality and Social Psychology* 103(5): 854–878. doi: 10.1037/a0029629

Grijalva, E., and L. Zhang. 2016. Narcissism and self-insight: A review and meta-analysis of narcissists' self-enhancement tendencies. *Personality and Social Psychology Bulletin* 42(1): 3-24. doi: 10.1177/0146167215611636

John, O.P., and R.W. Robins. 1994. Accuracy and bias in self-perception: Individual differences in self-enhancement and the role of narcissism. *Journal of Personality and*

Social Psychology 66(1): 206-219. doi: 10.1037/0022-3514.66.1.206

Kass, D.S., J.S. Kim, W. Zhu, and S. Erdley-Kass. 2021. The effects of self-other rating agreement on ethical behavior: An experimental investigation. *International Journal of Organization Theory & Behavior* 24(4): 273-288. doi: 10.1108/IJOTB-08-2020-0145

Kuenzi, M., M.E. Brown, D.M. Mayer, and M. Priesemuth. 2019. Supervisor-subordinate (dis)agreement on ethical leadership: An investigation of its antecedents and relationship to organizational deviance. *Business Ethics Quarterly* 29(1): 25-53. doi: 10.1017/beq.2018.14

McKee, R.A., Y. Lee, L. Atwater, and J. Antonakis. 2018. Effects of personality and gender on self-other agreement in ratings of leadership. *Journal of Occupational and Organizational Psychology* 91(2): 285-315. doi: 10.1111/joop.12209

Ou, A.Y., A.S. Tsui, A.J. Kinicki, D.A. Waldman, Z. Xiao, and L.J. Song. (2014). Humble chief executive officers' connections to top management team integration and middle managers' responses. *Administrative Science Quarterly* 59(1): 34-72. doi: 10.1177/0001839213520131

Owens, B.P., and D.R. Hekman. 2012. Modeling how to grow: An inductive examination of humble leader behaviors, contingencies, and outcomes. *Academy of Management Journal* 55(4): 787–818. doi: 10.5465/amj.2010.0441

Owens, B.P., and D.R. Hekman. 2015. How does leader humility influence team performance? Exploring the mechanisms of contagion and collective promotion focus. *Academy of Management Journal* 59(3): 1088-1111. doi: 10.5465/amj.2013.0660

Owens, B.P., M.D. Johnson, and T.R. Mitchell. 2013. Expressed humility in organizations: Implications for performance, teams, and leadership. *Organization Science* 24(5): 1517-1538. doi: 10.1287/orsc.1120.0795

Raskin, R., and C. Hall. 1979. A narcissistic personality inventory. *Psychological Reports* 45(2): 590. doi: 10.2466/pr0.1979.45.2.590

Rego, A., A.L. Melo, D.J. Bluhm, M.P. Cunha, and D.R. Júnior. 2020. Leader-expressed humility predicting team psychological safety: A personality dynamics lens. *Journal of Business Ethics* 174(2): 669-686. doi: 10.1007/s10551-020-04622-1

Ridge, J.W., and A. Ingram. 2017. Modesty in the top management team: Investor reaction and performance implications. *Journal of Management* 43(4): 1283-1306. doi: 10.1177/0149206314551796

Sosik, J.J. 2001. Self-other agreement on charismatic leadership: Relationships with work attitudes and managerial performance. *Group & Organization Management* 26(4): 484-511. doi: 10.1177/1059601101264005

Sosik, J.J., and L.E. Megerian. 1999. Understanding leader emotional intelligence and performance: The role of self-other agreement on transformational leadership perceptions. *Group & Organization Management* 24(3): 367-390. doi: 10.1177/1059601199243006

Swain, J.E. 2018. Effects of leader humility on the performance of virtual groups. *Journal of Leadership Studies* 12(1): 21-37. doi: 10.1002/jls.21552

Woo, S.E., C.S. Sims, D.E. Rupp, and A.M. Gibbons. 2008. Program engagement in response to feedback received in developmental assessment centers: Considering feedback favorability and self-assessor agreement. *Personnel Psychology*

61(4): 727-759. doi: 10.1111/j.1744-6570.2008.00129.x

Yammarino, F.J., and L.E. Atwater. 1993. Understanding self-perception accuracy: Implications for human resource management. *Human Resource Management* 32(2-3): 231-247. doi: 10.1002/hrm.3930320204

Yammarino, F.J., and L.E. Atwater. 1997. Do managers see themselves as other see them? Implications of self-other rating agreement for human resources management. *Organizational Dynamics* 25(4): 35-44. doi: 10.1016/S0090-2616(97)90035-8

Zenger, J., and J. Folkman. 2015. *What separates great HR leaders from the rest*. Retrieved August 12, 2021, from https://hbr.org/2015/08/what-separates-great-hr-leaders-from-the-rest