

# A Large Scale Survey of Motivation in Software Development and Analysis of its Validity - Reliability of Motivation Reports

Idan Amit                      Dror G. Feitelson  
idan.amit@mail.huji.ac.il      feit@cs.huji.ac.il  
Department of Computer Science  
The Hebrew University, Jerusalem 91904, Israel

October 4, 2024

## 1 Reliability of Motivation Reports

The limited reliability of motivation reports, a problem that we also cope with, was investigated in prior work. Using self-estimation in a survey might be a threat to the validity of the collected data. There might be biases due to ego defenses [2], the Dunning-Kruger effect [6], subjectivity, and different personal scales. Further, “research on self-esteem (Shavit & Shouval, 1980) [11] has demonstrated empirically that individuals resist lowering favorable self-perceptions” [4]. Previous work has tried to evaluate these difficulties.

Argyle [1] checked the reliability of self-estimation of happiness and showed it is related to peer and supervisory estimation. The Maslach Burnout Inventory validated self-estimation on burnout by comparison with the answers of a close person such as a spouse or a co-worker [7]. Judge et al. [5] also compared a person’s and significant other’s answers. For work answers “The average correlation between the self and significant-other reports, corrected for unreliability, was  $r = .68$ .”

Wigert and Harter investigated performance reviews, an area close to motivation [13]. They mention methodological difficulties when one tries to rely on supervisory estimation instead of self-estimation: individual supervisory ratings are a much less reliable measure of performance than objective measures [12], and 62% of the variance in ratings can be attributed to rater bias, while actual performance accounts for just 21% of the variance [10]. Yet, Tsui reports that an employee and his manager’s evaluation of effectiveness match [8].

Beatty et al. [3] also compare manager and employee’s appraisals. They found that there is agreement on medium performance and some disagreements on high and low performance. In a second usage there was higher agreement, though it was not clear if it was due to clarification of requirements or just better communication.

As prior work shows, there is a moderate agreement between self-reports and a close person’s report. This supports the self-reported answers validity yet warns that they are not perfectly accurate. In this study, we compare the same person’s answers to related questions, and the same person’s answers in the original and follow-up surveys, reaching a similar agreement level. We also note that despite all the above concerns, Scott et al. report that Facebook found that surveys are twice more accurate than predictive analytics in employee churn [9].

## References

- [1] ARGYLE, M. Do happy workers work harder? the effect of job satisfaction on job performance. In *How harmful is happiness? Consequences of enjoying life or not*, R. Veenhoven, Ed. Universitaire Pers, Rotterdam, The Netherlands, 1989, pp. 94–105.
- [2] BASSETT-JONES, N., AND C. LLOYD, G. Does herzberg’s motivation theory have staying power? *Journal of Management Development* 24 (Dec 2005), 929–943.
- [3] BEATTY, R. W., SCHNEIER, C. E., AND BEATTY, J. R. An empirical investigation of perceptions of ratee behavior frequency and ratee behavior change using behavioral expectation scales (BES). *Personnel Psychology* 30, 4 (1977), 647–658.
- [4] CAMPBELL, D. J., AND LEE, C. Self-appraisal in performance evaluation: Development versus evaluation. *The Academy of Management Review* 13, 2 (1988), 302–314.
- [5] JUDGE, T. A., LOCKE, E. A., DURHAM, C. C., AND KLUGER, A. N. Dispositional effects on job and life satisfaction: The role of core evaluations. *Journal of Applied Psychology* 83 (1998), 17–34.
- [6] KRUGER, J., AND DUNNING, D. Unskilled and unaware of it: How difficulties in recognizing one’s own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology* 77 (1999), 1121–1134.
- [7] MASLACH, C., JACKSON, S., AND LEITER, M. Maslach burnout inventory third edition. In *Evaluating Stress: A Book of Resources*, C. P. Zalaquett and R. J. Wood, Eds. Scarecrow Education, 1997, pp. 191–218.
- [8] S. TSUI, A. A role set analysis of managerial reputation. *Academy of Management Proceedings* 1982 (Aug 1982), 265–269.
- [9] SCOTT, J., O’ROURKE, E., AND GRANT, A. Employee surveys are still one of the best ways to measure engagement. *Harvard Business Review* (March 2018).

- [10] SCULLEN, S., MOUNT, M., AND GOFF, M. Understanding the latent structure of job performance ratings. *The Journal of Applied Psychology* 85 (Jan 2001), 956–70.
- [11] SHAVIT, H., AND SHOUVAL, R. Self-esteem and cognitive consistency effects on self-other evaluation. *Journal of Experimental Social Psychology* 16, 5 (1980), 417–425.
- [12] VISWESVARAN, C. V., ONES, D., AND SCHMIDT, F. Comparative analysis of reliability of job performance ratings. *Journal of Applied Psychology* 81 (Oct 1996), 557–574.
- [13] WIGERT, B., AND HARTER, J. Re-engineering performance management. *Gallup.com* (2019).