How to undo biased self-assessments

Marian Krajč, Andreas Ortmann, and Dmitry Ryvkin

Global Development Network / CERGE-EI Foundation

Regional Research Competition Project VIII+85, 2009

Non-technical summary

We report the results of two experiments (one field, one laboratory) through which we examined the impact of general information and specific information (feedback) on the quality of self-assessment ("calibration") in various tasks and feedback conditions. We find a strong positive effect of general information on calibration in the field experiment (Experiment 1); where in the case of exam predictions, students received direct feedback after the midterm exam. However, we observe improved calibration already before this information was revealed – this is most likely based on indirect feedback obtained from course work and interactions with classmates. In the laboratory experiment (Experiment 2) where we could better control specific information, the subjects who received specific information were calibrated better than those who did not. These results bear on the debate about the reality of cognitive illusions, and the question of how fast and far, and under what circumstances, the frequency of anomalies decays. At least for the allegedly well-established overconfidence phenomenon – one of the bones of contention in this debate, – it seems to take surprisingly little to reduce miscalibration quite dramatically.

In Experiment 2, subjects were comparatively well calibrated in absolute performance at all stages. This is mainly due to the design feature that, unlike in Experiment 1, subjects made predictions after performance and thus were familiar with the stimuli and could better assess their ability in performing the tasks. Our use of representative stimuli may have also contributed to this result. At the same time, calibration in percentiles in Experiment 2 is not better than in Experiment 1, indicating that, as expected, the timing of predictions does not matter for overconfidence.

Maybe unsurprisingly, in both experiments it is the unskilled who improve their calibration most. Thus, our results suggest that the unskilled may not be doomed to be unaware (if indeed they are).

In none of our experiments we observe anyone with predicted/estimated percentile rank in the worst 20% of the group. This is an interesting design problem potentially pertaining to all overconfidence studies that neither our study nor other studies so far have successfully addressed and that might be a reason for excessive overestimation of the unskilled.