ON IDENTIFYING HUMAN CAPITAL: FLAWED KNOWLEDGE LEADS TO FAULTY JUDGMENTS OF EXPERTISE BY INDIVIDUALS AND GROUPS

David Dunning

ABSTRACT

Purpose – *To thrive, any individual, organization, or society needs to separate true from false expertise. This chapter provides a selective review of research examining self and social judgments of human capital – that is, expertise, knowledge, and skill. In particular, it focuses on the problem of the "flawed evaluator": most people judging expertise often have flawed expertise themselves, and thus their assessments of self and others are imperfect in profound and systematic ways.*

Methodology/approach – *The review focuses mostly on empirical work specifically building on the "Dunning–Kruger effect" in self-perceptions of expertise (Kruger & Dunning, 1999). This selective review, thus, focuses on patterns of error in such judgments.* AU:1

Advances in Group Processes, Volume 32, 149–176

ISSN: 0882-6145/doi:10.1108/S0882-614520150000032006

Findings – *Because judges of expertise have flawed expertise themselves, they fail to recognize incompetence in themselves. Because of their flaws, most people also fail to recognize genius in other people and superior ideas.*

Practical implications – *The review suggests that organizations have trouble recognizing those exhibiting the highest levels of expertise in their midst. People in organizations also fail to identify the best advice and correct flawed ideas. Organizations may also rely on the "wisdom of crowds" strategy in situations in which that strategy actually misleads because too few people identify the best idea available.*

Keywords: Human capital; self-assessment; peer-assessment; Dunning–Kruger effect; advice; wisdom of crowds

The total wealth of an individual, organization, or nation does not include solely financial assets owned or the material resources that can be sold. Since the mid-twentieth century, economists have recognized that the skills, talents, and knowledge that individuals and groups possess comprise an important form of wealth or capital as well (Becker, 1964). For example, a licensed surgeon dropped into the middle of a city has skills and expertise that he or she can convert to greater prosperity than someone whose only skill is typing 25 words a minute. This *human capital*, which comprises a person's intellectual skills and technical knowledge, has increasingly become a focus of economists trying to explain the impact of schooling on the economy (Denison, 1962; Schultz, 1963), how the acquisition of skill influences wages (Mincer, 1974), and how countries develop advantages in trade (Findlay & Kierzkowski, 1983). The importance of human capital can be summed up by the estimate of economist Becker (2002) that 75% of the capital within the United States lies in the skills and knowledge of its citizens.

As such, for people, organizations, and societies to gain wealth and bolster well-being, they do well to develop their own human capital and learn how to exploit it. Central to this task is the capacity to *recognize* human capital – where it is present and where it is absent. The individual thrives if he or she can identify personal skills that can be utilized to best navigate the modern world, as well as weaknesses that should or must be improved upon. Organizations, from small groups to entire nations, succeed to the extent that they can identify or develop true experts within their midst who provide the best recommendations and leadership, while dismissing false authorities who mislead.

But herein lies the problem. How does one succeed in recognizing human capital? In economics, research suggests that people depend on such characteristics as years in schooling as a signal that one has acquired human capital (Denison, 1962; Schultz, 1963). But such indicators are not failsafe. A student may obtain a college degree but still suffer severe holes and deficits in important intellectual skills. After all, in handling a question about the central economic concept of *opportunity cost*, fewer than 22% (slightly less than chance) of nearly 200 professional economists attending an annual conference of the American Economic Association got the question right (Ferraro & Taylor, 2005). In a survey of roughly 2,300 students from 24 different universities in the United States, Arum and Roksa (2011) discovered that 36% of students displayed no improvement in their writing, critical thinking, and complex reasoning skills after four years of university study (Arum & Roksa, 2014).

Often, people or organizations go beyond social signals to construct their own assessment of skills and know-how, relying on "informal" methods of assessment based on intuition, commonsense theories, and homespun deliberation of what skill should look like. For example, Google famously asked such questions as *How much would you charge to wash all the windows in Seattle?* or *Why are manhole covers round?* (Moss, 2014) to gauge skills and brainpower in potential employees. However, it abandoned such questions in 2013 after empirical investigation showed the questions were completely uninformative of future performance (Moore, 2013).

In this chapter, I focus on informal, everyday assessments of knowledge and expertise, outlining some systematic problems that arise in such assessments, whether those evaluations be of self or other people. I argue that the recognition of human capital is not a straightforward task. If left to their own devices, constrained only to their own wits or erudition, both individuals and groups will make systematic errors in their judgments about who has know-how versus who knows not.

THE PROBLEM OF THE INFORMAL EVALUATOR

People will suffer difficulty when informally evaluating human capital because of one central observation. That contention is that the person judging expertise is usually, by definition, doing so under the shadow of his or her own inevitable incompetence. Except for the most trivial of tasks, each individual has gaps and flaws in his or her skills and knowledge. To be sure, some people possess gaps that are far wider than other people, as well as mistaken ideas that damage their judgments in occasionally flamboyant ways (Dunning, 2011; Dunning, Johnson, Ehrlinger, & Kruger, 2003; Ehrlinger, Johnson, Banner, Dunning, & Kruger, 2008; Kruger & Dunning, 1999). Others in contrast have knowledge and proficiency that far outstrip anything their peers might possess. But the central contention herein is that each individual sooner or later hits a ceiling of competence that lies lower than that describing total competence or perfection. As painter Salvador Dali was fond of saying, "Have no fear of perfection – you'll never reach it."

The inevitability of imperfections in one's own competence leads to one general implication in many areas of life. Imperfection not only impairs performance, it also impairs a person's ability to accurately *judge* performance. To the extent that people have gaps or defects in their expertise, they will not only perform imperfectly but will also make mistakes in their informal judgments of performance, whether it be by self or others. Imagine, for example, that you were asked to grade student performance on a classroom exam, but are asked to generate your own answer sheet to use to grade. To the extent that your personal answer sheet contains flaws and omissions, it will lead to you to giving grades that are flawed.

In many areas of life, such imperfection or incompetence thus leads to the individual to suffer a double curse: that person will not only produce flawed performance but will produce flawed judgments of performance as well. People will suffer this double curse because the skills or knowledge they need to produce a correct response are often the very same ones they need to judge the quality of that response. For example, the expertise needed to produce logically sound arguments is exactly the same knowledge needed to recognize whether a person has just made a logically sound argument. Generating a valid physics proof requires the same math skills needed to check the validity of the proof. Thus, the "skill set" needed to attain adequate performance on some intellectual or cognitive task is the exactly the same needed to accurately perform the *metacognitive* task of evaluating that performance, where metacognition refers to the evaluation of one's knowledge, reasoning, and learning (Metcalfe & Shimamura, 1994). In these cases, imperfection in tackling the cognitive task implies similar deficiencies in executing the metacognition one.

To be sure, there are some tasks that do not share this property. For example, running fast requires good legs and quick feet. But the skill set required to assess how quickly a person runs does not involve those legs and feet. Instead, it entails a stopwatch and enough attention to start and stop it appropriately. In this case, assessing running speed presents none of the problems suggested earlier. However, to the extent that performance and judgment of that performance converge to require the same expertise, people will be double cursed by their shortcomings and defects. Many tasks in life, falling in intellectual, social, or technical realms, present this doubly cursed challenge.

IMPLICATIONS FOR SELF-ASSESSMENT

Close to 20 years of research on self-assessment has shown that this double-curse analysis of expertise carries many important implications. Despite the Delphic admonition to "know thyself," contemporary psychological research suggests that people tend to have, at best, tenuous to meager insight into their competence, character, and personality (Dunning, 2005; Dunning, Heath, & Suls, 2004). To be sure, the perceptions people have of themselves tend to be correlated with the reality of their performance and behavior, but the relationship tends is far from perfect (Freund & Kasten, 2012; Hansford & Hattie, 1982; Harris & Schaubroeck, 1988; Zell & Krizan, 2014).

When it comes to judgments of expertise, a substantial portion of the disconnect between what people think about their skill and the reality of that skill when tested can be attributed to the "double curse." Consider the classic meta-analysis on self-evaluation published by Mabe and West (1982). When considering skills likely not inflicted with Dunning–Kruger issues, in that the competencies needed to judge skill are different from those necessary to produce good performance, one sees a higher correspondence between self-perception and reality. For example, judgments about physical skills tend to correlate (.47) with objective performance. However, when looking at skills more likely to be inflicted with the double curse, one finds smaller correspondence between self-perception and reality. Judgments of intelligence (.34) and of technical (.33), mechanical (.20), and medical-related (.17) skills, all tend to be only modestly related to actual skill. Judgments of social skills show little correspondence with reality, as evidenced in managerial (.04), job interview (.28), and interpersonal (.17) assessments.

Failing to Recognize Poor Self-Performance

But more direct tests of the double curse came from a comparison of those with low levels of skill to those with higher levels. Less skill in performance should also mean less accuracy in judgment. Thus, performers with low levels of competence should provide less accurate opinions of their own or anyone else's competence.

Ultimately, this would mean that low performers would have little insight into just how bad their performances are. Because they, like everyone else, choose what they think is the most reasonable course of action, they should show confidence in the choices they make. They, unfortunately, do not have the expertise to recognize the errors they make, although other people can. They also cannot recognize the superiority of other people's decisions when presented with them. As a consequence, they remain blithely unaware of the depths of their incompetence. This state of affairs among the incompetent has since become known in popular culture as the Dunning–Kruger effect (e.g., ~~Kay & Shipman, 2014;~~ Morris, 2010), and AU:2 many studies have demonstrated its existence.

An example of data from a typical Dunning–Kruger effect study can be found in Fig. 1 (Dunning et al., 2003). This figure represents the relationship between college student perceptions of how well they have performed on a course exam they have just completed and how well they have actually performed. The responses are real; these data are taken from a large-lecture psychology class at Cornell University. In the figure, students have been sorted into four groups based on their actual test performance, from the bottom 25% of performers to the top 25%. Students' perceptions of "their mastery of course material" and their "performance on this specific test" are tracked as a function of their performance group. Students are asked to report the percentile they think their performance falls in.

The figure shows three (actually, four) major findings. First, students on average overrate their performance. By definition, the average performance falls on the 50th percentile, the point at which as many peers outperform the student as the student outperforms. All groups, even the bottom performers, believe they outperform a majority of their peers, with the average self-rating in each group falling at least in the 60th percentile or above. In a sense, this is not a surprise. Psychological research consistently shows that people overestimate their performance, skill, status in the world, and their future prospects (for reviews, see Dunning, 2005; Dunning et al., 2004). Second, there is a significant but shallow relationship between perception of performance and reality, with top performers rating their achievement

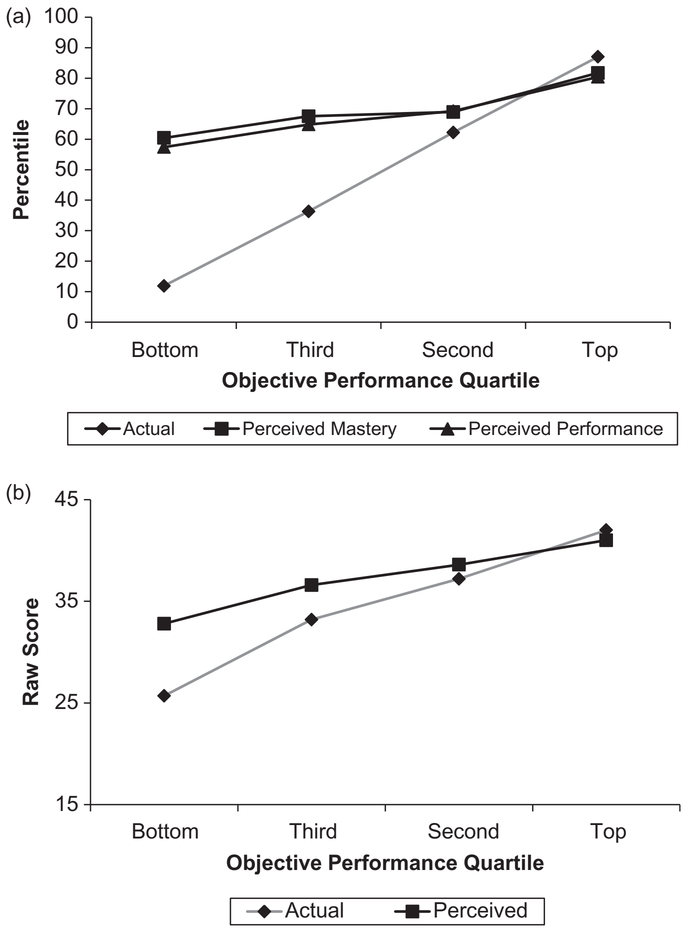

Fig. 1 Perceived Performance as a Function of Objective Performance on a Class Exam. (a) Presents Percentile Ratings for Perceived Mastery of Course Material and Performance on the Exam. (b) Presents Perceived Raw Score on the Exam (Out of 45 points). *Source*: From Dunning et al. (2003). Adapted with permission.

slightly above bottom performers. Again, this pattern echoes past research on self-insight, or the lack thereof, into skill and expertise (Dunning, 2005; Dunning et al., 2004; Mabe & West, 1982).

It is the third finding that is the most important. Focusing on bottom performers, one sees that they rate their knowledge and performance rather highly, in the 60th percentile, although their actual performance lies much lower, at the 12th percentile. Similar analyses focusing on estimates of raw test scores show a similar pattern, with bottom performers overestimating their raw performance by 20–40%.

Not only found in the classroom, we have found this pattern among students taking logical reasoning or grammar quizzes, or assessing their sense of humor (Kruger & Dunning, 1999). We have seen it among debate teams taking part in college tournaments (Ehrlinger et al., 2008), and other researchers have seen the same among participants in chess tournaments and weekly bridge competitions (Simons, 2013), as well as among participants at a trap-and-skeet competition doing badly on a quiz concerning firearm care and safety (Ehrlinger et al., 2008). It has been observed among medical lab technicians queried about aspects of their job, international students seeking pharmacy licensure in Canada, and obstetrics/gynecology interns completing their rotations (Austin, Gregory, & Galli, 2008). In each case, poor performers show little insight into how poorly they are really doing (for a recent review, see Dunning, 2011).

This lack of insight appears to be an honest one. Offering to pay people up to $100 for accurate self-assessments does nothing to enhance the quality of their self-judgments or to make poor performers more negative about their performer (Ehrlinger et al., 2008). Being exposed to how other people approach a task or test – thus, seeing how other people approach the same tasks differently – does nothing to make poor performers revise their flattering and inappropriate self-estimates of skill (Hodges, Regehr, & Martin, 2001; Kruger & Dunning, 1999)

Implications for Behavior

These misperceptions also influence behavioral choices. In a clever study, Ferraro (2006) presented 17 college students with two answer sheets to the 10 hardest questions students had faced in a recent economics exam. One sheet contained the answers the student had provided to those questions on the exam. The other contained answers that were all correct. Students were asked to choose the answer sheet they thought represented the best performance on those questions. For every question answered correctly on the sheet they selected, they would receive one extra credit point toward their final exam grade. Even though students themselves in reality got more than half the questions wrong, 13 of 17 (over 75%) chose the answer sheet that represented their own responses over the one with only correct responses.

In another ingenious follow-up, Ferraro (2010) presented students with "insurance policies" that hedged against any bad performance they might encounter in a classroom exam. Policy A cost 10 exam points (which would be really deducted from the exam), but it would add 20 points to student scores if they finished in the bottom half of the class. Policy B cost only 2 exam points, but would pay only 4 points if the student's score fell between the 50th and 75th percentile. Given these terms, 50% of students would have "bought" Policy A and 25% Policy B had they had perfect insight into their exam performance. However, only 33% of students bought a policy, suggesting that many expected to perform better than the 75th percentile. And of the two policies, Policy B turned out to be more popular than A, which again suggested that students expected their performance to be more likely to fall between the 50th and 75th percentile than to fall in the bottom 50%. In fact, of 19 students who had fallen in the 50% on every previous exam in the class, only four bought Policy A, while six bought Policy B.

CRITICS OF THE FRAMEWORK

To be sure, these findings and our analysis of them are not without critics. Other researchers have asserted that the Dunning–Kruger pattern of self-error is mere statistical artifact. For example, some researchers have argued that the pattern is simply a regression-to-the-mean effect (Ackerman, Beier, & Bowen, 2002; Burson, Larrick, & Klayman, 2006; Krueger & Mueller, 2002). Simply because of measurement error, perceptions of performance will fail to correlate perfectly with actual performance. This dissociation due to measurement error will cause poor performers to overestimate their performance and top performers to underestimate theirs, the pattern found, for example, in Fig. 1. In response, we have conducted studies in which we estimate and correct for measurement error, asking what the perception/reality link would look like if we had perfectly reliable instruments assessing performance and perception. We find that such a procedure reduces our pattern of self-judgment errors only trivially (Ehrlinger et al., 2008; Kruger & Dunning, 2002).

Other researchers, working from assumptions about the performance distribution underlying our data, have claimed that poor performers face a more difficult job estimating their percentile ranking simply because there are more poor performers than top performers at the rarefied level of expert achievement (Krajč & Ortmann, 2008). However, when we look, we fail to find evidence for the critical assumptions these critics make to support their analysis, nor much adjustment to our findings if we go ahead and correct for these assumptions anyway (Schlösser, Dunning, Johnson, & Kruger, 2013).

MANAGING THE INCOMPETENT

The implications of self-errors inspired by the Dunning–Kruger effect carry obvious implications for the individual making them. But these errors also produce challenges for others or organizations who must deal with the individuals making these errors.

The Paradox of Advice

Some of these implications will be obvious to anyone involved in management. To the extent that a manager must counsel and mentor others, his or her success may very well depend on the competence and expertise of the underling receiving the counsel. All academic advisors know this: often, it is the students who need advice who are the ones least likely to know it and to show up at their academic advisor's office. Even if they do, they often are the ones who resist or discard any advice they receive.

In short, incompetence leads to many paradoxes when it comes to advice. Incompetence can make people unaware that they need advice. In addition, to weigh advice appropriately, one must already have enough expertise to assess the worth of the advice we receive. If one does not have that expertise, then one runs into trouble separating good advice from bad.

We have recently demonstrated this in the advice paradigm commonly used in the organizational psychology literature (Bonaccio & Dalal, 2006). In that literature, participants are asked to estimate some value, such as how tall Mt. Everest is or what year the Wright Brothers flew their first airplane at Kitty Hawk. After giving their estimate, participants are provided the estimate of another individual as advice and asked if they wish to revise their initial estimate. Usually, people revise their estimate somewhat toward the one provided, giving their own initial estimate more weight in their final conclusion (Bonaccio & Dalal, 2006; Yaniv & Kleinberger, 2000). This turns out to be a mistake in the long run: if they moved to a point equally between their own initial estimate and the "advice" they were given, they would be significantly more accurate.

How does weight given to advice vary according to a person's own expertise? Do people know when they are right versus grossly wrong? And, more important, can they differentiate good advice from bad? I conducted a study (Dunning, 2014b) in which American participants were asked to estimate the years in which 12 different historical events took place (e.g., the founding of CNN, the end of the Spanish-American war, and Hawaii became a state). Then, for six of the events, participants were given a previous respondent's advice – which happened to be pretty good, namely the actual years those events took place. For the other six, respondents were given bad advice representing some of the worst estimates given in a previous group. I also assessed how much expertise respondents revealed in their initial estimates. Were they, on average, close to the actual years the events took place or were they further away.

How well participants weighed the advice they were given depended on their own expertise, as shown in Fig. 2. For good advice, both knowledgeable (+1 SD from the mean in accuracy) and unknowledgeable (−1 SD from the mean of accuracy) gave equal and substantive weight to the advice, moving over halfway from their initial estimate toward the provided estimate in their final answer (51% and 59%, respectively). However, when it came to bad advice, knowledgeable and unknowledgeable respondents differed. Knowledgeable respondents hardly budged from their initial estimates, moving only 6% toward the bad advice. Unknowledgeable respondents, in contrast, gave the bad advice much weight, moving 58%

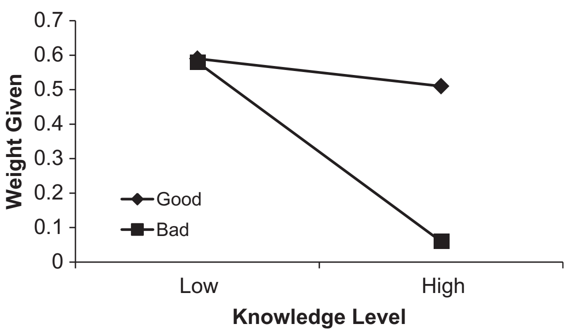

Fig. 2 Weight Given to Good and Bad Advice as a Function of Participant Knowledge Level. *Source:* From Dunning (2014).

toward it from their original estimate, nearly just as much as they moved toward good advice.

The Fate of Feedback

In addition, feedback *per se* is not necessarily an effective means to rid people of their inability to see their incompetence. Students taking multiple exams in a class do not become more accurate in predicting their exam scores as they gain experience (Ferraro, 2010; Schlösser et al., 2013). If anyone does gain accuracy, it tends to be the more competent students. Students performing the worst do not alter their predictions about exam performance to become more accurate (Hacker, Bol, Horgan, & Rakow, 2000).

The reasons for this resistance to gaining insight into incompetence may be many. When predicting future performance, people give less weight to their past performance than they do to how well they aspire to perform (Helzer & Dunning, 2012) – even though they give heavy weight to past performance in their predictions of other people. For example, in one study, we gave students an opportunity to predict the upcoming exam performance of one of their peers. They could win up to $5 for accurate predictions. Given the opportunity, participants wanted to learn about their peer's past performance in the class rather than about the performance they aspired to. However, we then played a second prediction game with these same participants. Some peer was to predict how the participant himself or herself was to perform on an upcoming exam, with accurate predictions earning the participant up to $5. What information did the participant now want to give the peer making the prediction? Most participants chose to give information about the performance they aspired to in the future rather than information about their past performances (Helzer & Dunning, 2012).

In addition, people resist negative feedback, sometimes with paradoxical results. In a telling study, we gave business students a chance to buy a self-improvement book on emotional intelligence at a discount after learning about the concept in their class and taking a test of their own emotional intelligence. The feedback we gave participants had a strong impact on whether they wanted the book, but not necessarily in the way one would expect. Over 65% of top performers bought the book, but only 20% of poor performers did likewise. Ironically, giving performance feedback inspired self-improvement only among the best performers and not among those who arguably needed it the most (Sheldon, Dunning, & Ames, 2014).

Additionally, people resist negative feedback because they can be quite creative and flexible about finding routes that allow them to dismiss negative feedback. In another emotional intelligence study, we asked one participant group about whether emotional intelligence was relevant to their lives before giving them feedback about their performance, thus "fixing" their impressions of relevance before receiving any feedback. Students performing badly in this group reacted by disparaging the accuracy of the test, which was then connected to a reluctance to engage in self-improvement. However, with another participant group, we asked them whether the test was accurate before we gave them feedback, thus fixing their opinions of test accuracy. Those in this group performing badly reacted by deriding the relevance of emotional intelligence for their lives, which was then connected, again, to a resistance to self-improvement emotional intelligence skills (Sheldon et al., 2014).

Only one intervention seems to work to make poor performers see the error of their ways, and it is a paradox: training poor performers to become competent. Now skilled, they recognize their previous responses as errors, and so become more modest in rating their skill after having, if anything, been trained them to be more proficient at that skill (Kruger & Dunning, 1999). The irony of competence is to be more skilled at seeing examples of one's own past incompetence; one can see more accurately the unskilled individual that one no longer is.

THE BURDENS OF TOP PERFORMERS

At the level of the collective or group, some of the burden imposed by the Dunning–Kruger framework fall disproportionately on some of its members more than others. In particular, top performers do not escape being affected by the imperfect expertise possessed by more typical members of the group.

Failing to See One's Uniqueness

Some of this burden is revealed by looking back at Fig. 1. Much like bottom performers, top performers also tend to misjudge the knowledge – but they do so in a way quite different from their less knowledgeable peers. In the figure, one sees that top performers tend to underestimate their performances relative to their peers – not seeing how unique or special their performance level is. They do so because top performers suffer from a "curse of knowledge" (Camerer, Loewenstein, & Weber, 1989; Nickerson, Baddeley, & Freeman, 1987). Namely they infer that other people must have the same level or similar levels of knowledge as themselves. If they know it, others must too. As a consequence, their errors in the figure come about via a very different psychology from that of poor performers. Top performers, unlike poor performers, assess their own work accurately in an objective sense. They have a more correct sense of what their correct responses are and what they likely get wrong (see Ehrlinger et al., 2008). Where they err is in their assessments of other people, overestimating what their peers know.

We have documented this dynamic by collecting information of what top performers think of their peers and then correcting for it statistically. Such a procedure removes a good deal of the overall judgmental errors that top performers make, although it does nothing for poor performers (Ehrlinger et al., 2008), who are mainly wrong about themselves. Additionally, exposing top performers to how their peers approach a task tends to "clue them in" about low levels of knowledge among their peers – and thus leads them to recognize just how unique or "special" their own performances and achievements are (Kruger & Dunning, 1999; see also Hodges et al., 2001).

Genius as Unrecognized

However, failing to recognize their own specialness is not the only burden that top performers suffer. As we have moved our research on the Dunning–Kruger framework from the individual to the collective level, we soon came to an important insight about how top performers are likely regarded by others. Consider a group assessing the expertise of the individuals within its midst. The group, as a whole, likely has the skill level to identify poor performers and their deviant mistakes. But consider what happens when the target of judgment becomes more competent. At each rise in competence, fewer in the group have the expertise to spot the true level of competence displayed by the target.

Ultimately, at the highest level of competence, the expertise of the target outstrips that of most in the group. Those individuals in the group likely fail to have the expertise to understand just how competent this high performer is. Thus, when his or her judgments deviate from these less competent peers, his or his judgments will be seen as potential errors rather than as genius. It is as Jonathan Swift once observed, "When a true genius appears, you can know him by his sign: that all the dunces are in a confederacy against him."

In short, genius will hide in plain sight. The group on average may have the expertise to spot any poor performers within it. However, the group's expertise on average will be imperfect, and thus it will largely fail to recognize individuals with competence closer to perfection. This assertion that genius hides in plain sight fits well with studies that showing that groups often have trouble identifying those performing the best among its midst (Littlepage, Schmidt, Whisler, & Frost, 1995; Miner, 1984).

It is also consistent with more recent data we have collected expressly to test the hypothesis. In those studies, we asked respondents to take tests of logical reasoning, financial literacy, or intuitive physics. We then give them tests filled out by previous participants, representing quite a range from poor to excellent performance, asking respondents to estimate how well each of these previous participants has done. Fig. 3 presents results from a typical study, depicting the average judgments of 37 respondents examining targets who scored 4, 8, 12, 16, or a perfect 20 on a test of logical reasoning (with average performance for the respondents hovering around 12). Half of the participants were given financial incentives (up to US$50) for accurate judgments; the other half heard no mention of incentives (Dunning & Cone, 2014).

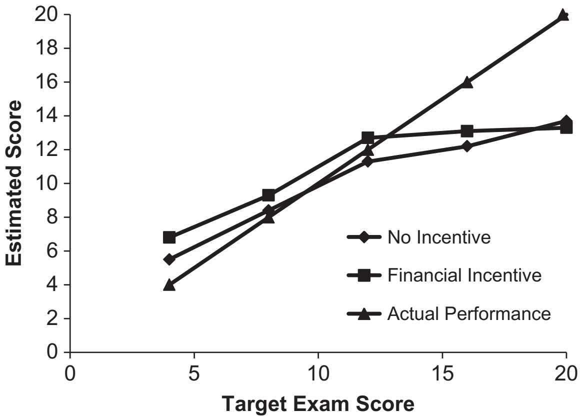

Fig. 3 Average Estimate of Target Performance of that Performance.
Source: From Dunning and Cone (2014).

As seen in the figure, the group is largely accurate in its judgments of poor-performing targets. However, as target performance increases, estimates and their accuracy begins to fall off. In fact, the performance of the top performer is underestimated by some 35% – with the group seeing this performance as barely above average. And, as seen in the figure, financial incentives had no impact on the group's accuracy. In addition, although not shown in the figure, higher performing respondents came closer to recognizing the truly excellent performance of the top performer relative to those doing more poorly (Dunning & Cone, 2014).

We have replicated this finding of genius hiding in plain sight in multiple ways. In one study, we ask participants which peer they would go to for financial advice after seeing a set of peer performances on a financial literacy test. A greater proportion of participants spotted the person they should most avoid (i.e., the bottom performer; 43.2%) than identified the person they should most seek out (i.e., the top performer; 29.6%), a difference reaching statistical significance (Dunning & Cone, 2014).

In addition, in another study, we asked participants to compare themselves to each individual target and bet whether their score on a logical reasoning test beat was beaten by or tied the target's score. They would win an additional $1 if their prediction was right. With the bottom performer, participants were largely accurate in their bets. Of the 101 respondents, 84 thought their score would beat that of the target, with additional 7 participants claiming ties. These predictions came quite close to reality (i.e., 89 wins, 6 ties). However, when comparing themselves to the top performer, respondents grossly overestimated themselves. Of the 101, 27 thought their score would beat that of the top performer, with an additional 27 claiming a tie. In reality, only one participant's performance managed to tie that of the top performer, with the rest losing (Dunning, 2014d).

Leading Back to Self-Flattering Social Comparisons

In a sense, the fact that genius hides in plain sight provides an intriguing explanation for an oft-documented phenomenon in self-psychology. That finding is that people on average think of themselves as anything but average. Among college students, 70% believe themselves to be "above average" in leadership ability, but only 2% see themselves as below average (College Board, 1976–1977). A full 94% of university professors state that they do above-average work relative to their peers (Cross, 1977). In one software development firm, a full 32% of engineers thought their skill level put them in the top 5% of employees; in another firm, the figure was 42% (Zenger, 1992). This above-average bias in self-assessment is ubiquitous (for a review, see Dunning et al., 2004). Ironically, people even claim above-average abilities in providing accurate self-evaluations untarnished by bias and distortion (Friedrich, 1996; Pronin, Lin, & Ross, 2002).

The inability to recognize talent among top-performing peers, coupled with a true capacity to spot poor performers, would lead to these inflated self-estimates, honestly believed. If accuracy is lop-sided, with people accurately spotting those they outperform but underestimating those who outperform them, then people would have sincere but mistaken evidence about how their expertise compares against that of their peers.

Again, we have recent evidence of this dynamic (Dunning, 2014d). In it, I asked participants to complete a reasoning test in which participants, on average, thought they scored 6 out of 10 items right, a significant but slight overestimate. Subsequently, for half of participants, I showed them tests as filled out by five poor performers who averaged a score of 3.4. For the other half, I showed them tests filled out by five superior performers who attained an 8.6 score on average. I asked all participants to estimate how well these other performers had done. Participants moderately overestimated the performance of the poor group by 1.2 items on average, but solidly underestimated the performance of the superior group by 2.0 items.

In short, participants accurately saw themselves as outperforming the poor performers, even though they moderately overestimated those performers. However, due to a mixture of overestimating themselves and underestimating superior performers, they failed to see just how much they came up short in comparison to those top-line performers. This was evident in assessments in which they directly compared their own performance to the set of targets they had been given. Participants rated themselves superior on average to the poor-performing group – an accurate result. However, they rated their own performance roughly equal to that of the superior performers, a grossly mistaken impression, but one that followed directly from their misestimates of self- and peer performances. In sum, blindness to performances better than one's own can lead to self-flattering social comparisons not supported by actual evidence.

Resistance to Correct Beliefs in the Marketplace of Ideas

What is true of top-performing people may also be true of very smart ideas. They, too, suffer burdens due to the Dunning–Kruger framework and might prove too smart for their own good. The collective may not share the expertise or "intellectual scaffolding" to understand their worth. As computer pioneer Howard Aiken put it, "don't worry about people stealing your ideas; if your ideas are any good, you'll have to ram them down people's throats." Or, as English writer Aldous Huxley observed, "the vast majority of human beings dislike and even dread all notions with which they are not familiar; hence it comes about that at their first appearance innovators have always been derided as fools and madmen." Smart ideas, especially those of the highest IQ, run the risk of being doubted, unheeded, or discarded in favor of notions that still carry some worth, but not that of the "high IQ" idea.

We have begun to conduct research to see if ideas might be too smart to be truly recognized by the collective. Central to this idea is the notion, taken from the literature on the marketplace of ideas, of the "habitat" in which ideas must live and survive. Some ideas have a friendlier habitat, in that people possess beliefs or knowledge that "hook into" those ideas more easily and validate them, thus supporting people's adoption of them. For example, many more recent stories about witches have arisen than stories about trolls because people already have a rich set of associations and stereotypes concerning the former then they do the latter. People are cognitively prepared to understand and remember any story about witches more than they about trolls, and thus stories about witches "win" in the marketplace of ideas (Berger & Heath, 2005). High IQ ideas, by definition, may exist in harsher habitats, in that people typically fail to have the expertise or intellectual scaffolding necessary to recognize their worth or even remember them accurately, relative to ideas understood by a larger proportion of the population.

I have already demonstrated this in an initial study (Dunning, 2014a), asking people how likely they are to pass along various pieces of information they have heard before to a friend who is interested. For example, participants are asked how likely it is that they will pass along the answer "the Sahara" if their friend wants to know what is the largest desert in the world, or "Pluto" if they friend wants to know which planet was discovered last. There is one trick to the study, however. For some participants, they are asked if they will pass along answers that turn out to be popular, but which actually turn out to be wrong, such as the Sahara and Pluto. The correct answers to those two questions are the Antarctic and Neptune (remember, Pluto has been decommissioned as a planet as of 2005).

However, when asked the likelihood of passing along the answers, participants reported being more willing, and more disposed to vouch as true, popular but wrong answers over unpopular but right ones. In a more telling example, participants stated they were more likely to pass along information to a friend worried about chronic high blood pressure that the disease producing symptoms of dizziness and headaches (which is inaccurate) over information that the condition had no external symptoms (which is correct) (Dunning, 2014a).

Limits to Correcting Belief

Correct ideas may also suffer resistance in another way if they fall too outside the box of people's intuitions. Given advice about these ideas, people may give them weight, but with limits. This was demonstrated in another study about a counterintuitive statistic – that the top 20% wealthiest of Americans possess roughly 89% of the total wealth in the country. In an experiment in which people were asked to estimate the percentage of wealth owned by that top 20%, respondents on average initially estimated that the figure was roughly 67%, as shown in Table 1. We then exposed participants to an estimate that had been given by a previous respondent, from 30% to 90% (the correct answer) by intervals of 15%, and asked if they wished to revise their estimates.

Participants showed some sensitivity to the accuracy of the estimates they were provided. Those exposed to lower, but inaccurate, estimates of 45% and 30% declined to lower their estimates. In the groups given 75% and 90% as an advisor's estimate, figures closer to the truth, both groups revised their estimates upward to a statistically significant degree. But there

Table 1. Impact of Other Person's Estimate on One's Own Estimate of Proportion of Wealth Held by Top 20% of the Wealthiest in the United States.

Other's Estimate (%)	Own Estimate (%)		*p*
	Initial	Final	
90[a]	64.1	73.2	<.001
75	67.9	73.3	<.001
60	68.3	68.5	ns
45	66.7	67.1	ns
30	66.5	63.9	ns

[a]Correct figure.

was a limit to how much participants were willing to move their estimates toward accuracy. Told that another participant thought it was 90%, participants moved their estimates – all the way to 73% – only about 35% of the way toward the correct answer – and roughly only the same as they did when told incorrectly that the figure was 75% (Dunning, 2014a).

Taken together, this work echoed past research on advice taking and classic work on persuasion. The good news is that respondents are more responsive to accurate advice than to inaccurate counsel (Gardner & Berry, 1995; Yaniv & Kleinberger, 2000). However, there are limits to how much people will revise their impressions. And if even accurate information lies too far outside their initial opinions and intuitions, people will discount the amount of weight they give it and the degree to which they will revise their beliefs (Bochner & Insko, 1966; Harries, Yaniv, & Harvey, 2004; Sherif & Hovland, 1961).

The Curtailed Wisdom of Crowds

The Dunning–Kruger framework also carries implications for reaching more accurate conclusions via judgment aggregation. A long history of scholarship has shown that simply aggregating the judgments of different people usually results in an aggregate judgment that is more accurate than even the most expert members of the group, or which can be derived by any more complex procedure (Surowiecki, 2004). That is, the quickest and most efficient way toward wisdom is merely to take the average of the collective wisdom of all the individuals in a group. Aggregating multiple judgments retains the component of "truth" contained in each individual's judgment while canceling out any random errors in individual judgments varying from the truth. In the classic example demonstrating this *wisdom of crowds*, the average of guesses of 800 attendees at a local livestock fair of the weight of one particular ox fell only one pound over the ox's true weight (Galton, 1907).

This "wisdom of crowds" notion is also used as a prime justification for democratic forms of government. The argument is that the votes of individuals might contain a good deal of error and misinformation, but by aggregating all those imperfect votes, one will arrive at a result that cancels individual errors while retaining any kernel of wisdom contained in each. Thus, democratic elections, as long as individual voters' choices contain some wisdom and there are enough votes to cancel out individual voter error, are likely to arrive at the best decisions in a nation's interests. This logic is encapsulated in Lord Condorcet's Juror Theorem, which at its core extols the wisdom of *statisticized* groups created by aggregating the democratic votes of its members (Austen-Smith & Banks, 1996; List & Goodin, 2001).

The theoretical framework here, however, suggests a limit to this logic, suggesting instead that statisticized groups will often likely arrive at moderately accurate judgments while rejecting the best or most accurate judgment available. This occurs because a majority of individuals will likely reason their way to a conclusion that is pretty good, but not better but more deviant ideas. That is, if only a minority of individuals has the intellectual skills necessary to arrive at rare but more accurate answers, their influence will wane as aggregation increases. Aggregation instead will ratify a more mediocre conclusion rather than the best one.

In my lab, we have run a series of simulations that demonstrate that aggregating judgments often throws away the best ideas (Dunning & Covington, 2014). Consider this classic example of a *disjunction problem* (Topla~~c~~k & Stanovich, 2002):

> Jack is looking at Ann but Ann is looking at George. Jack is married by George is not. Is a married person looking at an unmarried person? Yes, no, or it cannot be determined?

In a survey of 204 individuals, 156 stated that the answer could not be determined, but the correct answer is "yes," which only 45 respondents indicated: if Ann is unmarried, then a married Jack is looking at an unmarried person. If she instead is married, then once again a married person is looking at an unmarried one.

In all, 22% of respondents chose the right answer, and aggregating the judgments of all participants does not improve that figure. Fig. 4 shows the results of aggregating judgments of 3, 5, 8, 10, and 15 participants to determine a plurality answer. We ran 1,000 aggregation simulations at each group size, and found that the group returned the correct answer less frequently as its size increased, with less than 1% of groups of size 15 returning the right answer. In short, aggregating judgment tended to threw away the correct answer when that answer was a minority position in the group. Similar simulations show a similar effect in aggregation of the judgments of chess players asked to select the best move in various scenarios. Aggregating those judgments prompts bad moves to be rejected, but it also promotes the rejection of more esoteric but superior moves favored by grandmasters (Dunning & Covington, 2014). Specifically, if fewer than 25% of the chess players choose the best move individually, that move will

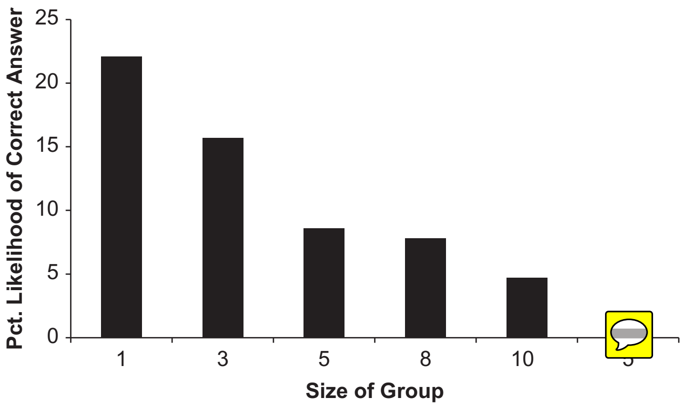

Fig. 4 Likelihood of Arriving at Correct Answer for the "Marriage Problem" as a Function of Group Size. *Source*: From Dunning and Covington (2014). AU:4

be rejected more frequently in favor of a more mediocre move by any aggregation of multiple chess players.

Nagel (2010) has extended this logic to a formal analysis of democratic elections in which people can accurately judge the leadership ability of those they outperform but provide only random judgments regarding the skill of people who outperform them. Under these conditions, a series of simulations showed that both bad and very good leaders were rejected in democratic elections, leaving the group to be ultimately led by leaders who had somewhat above average but hardly outstanding leadership skill.

To wit, statisticized crowds prove wise enough to reject bad options, but not wise enough to adopt the very best options available. There are limits to its recognition of human capital.

CHALLENGES FOR FUTURE RESEARCH AND BEYOND

At the beginning of this chapter, I noted the importance of human capital for well-being and growth, ranging all the way from the individual to organization to society or nation. Both individuals and groups have to identify sources of expertise that may lie in their midst, or know when such expertise must be sought out or developed.

I then developed one problematic aspect of the human condition when it comes to the identification and assessment of human capital: often, those making the assessments do so with imperfect expertise. As a consequence, it is in inevitable that their assessments of human capital, both in self and others, will be flawed, sometimes dramatically so. I outlined various consequences that fall out from that analysis. When it comes to self-assessment, people will have little insight into their intellectual and social deficits. When it comes to judging other people, communities will largely succeed in identifying poor performers in their midst but suffer difficulties in identifying the most able. This will lead them to miss the appearance of genius among them, embodied both in terms of individuals and ideas. Usual methods for improving the identification genius, such as relying on the wisdom of the crowd, may not work.

All these difficulties present challenges, both for future research and for practitioners. What does one do when the information identification of human capital is fallible? One obvious notion would be to move to more formal methods of assessment – more objective tests of skill and expertise. However, such tests may do a better job at assessing some well-defined skills, technical skills, for example, but do a less precise job at assessing no less important but less-definable skills, such as professionalism or social skills (Veloski, Fields, Boex, & Blank, 2005).

In addition, one might pay more attention to past performance as an indicator of expertise going forward (Mannes, Soll, & Larrick, 2014; Mellers, Ungar, Baron, & Tetlock, 2014). However, in doing so, one must make sure that one is in a domain in which superior performance during one period is predictive of future ones, or will continue if important conditions change. There are significant human endeavors in which such predictability does not exist, such as in forecasts of financial markets, in which yesterday's soothsayer typically becomes today's guru only to become tomorrow's source of disillusionment (Malkiel, 2012).

Thus, moving forward, it appears that a central challenge facing both researchers and practitioners is identifying expertise where ground truth is not necessarily known. One way to do so is to strive for further expertise, in that the problems associated with the Dunning–Kruger framework evaporate as ignorance is dispelled (Dunning, 2011; Kruger & Dunning, 1999). One might also exploit the ability of crowds to identify poor performers, eliminating their contribution to group judgments or statisticized opinions before starting the aggregation.

Whatever method may be yet invented to aid in the identification of human capital, we can presume that its benefits will be large and far-reaching. As Alfred Marshall (1890) stated in his *Principles of economics*, "the most valuable of all capital is that invested in human beings." What I argue herein is not only a concurrence of Marshall's assertion, but that such investment is not only valuable for its direct benefits. Such investment provides crucial guidance, for both individual and collective, in identifying where that human capital may lay or instead be in need.

ACKNOWLEDGMENT

Preparation of this chapter, and conduct of much of the research reported herein, was supported by a THRIVE center grant entitled "Cognitive habits of intellectual humility," funded by the Templeton Foundation.

REFERENCES

Ackerman, P. L., Beier, M. B., & Bowen, K. R. (2002). What we really know about our abilities and our knowledge. *Personality and Individual Differences*, *34*, 587–605.

Arum, R., & Roksa, J. (2011). *Academically adrift: Limited learning on college campuses*. Chicago, IL: University of Chicago Press.

Arum, R., & Roksa, J. (2014). *Aspiring adults adrift: Tentative transitions of college graduates*. Chicago, IL: University of Chicago Press.

Austen-Smith, D., & Banks, J. S. (1996). Information aggregation, rationality, and the Condorcet Jury Theorem. *American Political Science Review*, *90*, 34–45.

Austin, Z., Gregory, P. A. M., & Galli, M. (2008). "I just don't know what I'm supposed to know": Evaluating self-assessment skills of international pharmacy graduates in Canada. *Administrative Pharmacy*, *4*, 115–124.

Becker, G. S. (1964). *Human capital: A theoretical and empirical analysis*. Chicago, IL: University of Chicago Press.

Becker, G. S. (2002). The age of human capital. In E. P. Lazear (Ed.) *Education in the twenty-first century* (pp. 3–8). Palo Alto, CA: Hoover Institution Press.

Berger, J. A., & Heath, C. (2005). Idea habitats: How the prevalence of environmental cues influences the success of ideas. *Cognitive Science*, *29*, 195–221.

Bochner, S., & Insko, C. A. (1966). Communicator discrepancy, source credibility, and opinion change. *Journal of Personality and Social Psychology*, *4*, 614–621.

Bonaccio, S., & Dalal, R. S. (2006). Advice taking and decision-making: An integrative literature review, and implications for the organizational sciences. *Organizational Behavior and Human Decision Processes*, *101*, 127–151.

Burson, K. A., Larrick, R. P., & Klayman, J. (2006). Skilled or unskilled, but still unaware of it: How perceptions of difficulty drive miscalibration in relative comparisons. *Journal of Personality and Social Psychology*, *90*, 60–77.

Camerer, C., Loewenstein, G., & Weber, M. (1989). The curse of knowledge in economic settings: An experimental analysis. *Journal of Political Economy*, *97*, 1232–1254.

Caputo, D. D., & Dunning, D. (2005). What you don't know: The role played by errors of omission in imperfect self-assessments. *Journal of Experimental Social Psychology*, *41*, 488–505.

College Board. (1976–1977). *Student descriptive questionnaire*. Princeton, NJ: Educational Testing Service.

Cross, P. (1977). Not can but *will* college teaching be improved? *New Directions for Higher Education*, *17*, 1–15.

Denison, E. F. (1962). *Sources of economic growth in the United States*. Washington, DC: Committee for Economic Development.

Dunning, D. (2005). *Self-insight: Roadblocks and detours on the path to knowing thyself*. New York, NY: Psychology Press.

Dunning, D. (2011). The Dunning-Kruger effect: On being ignorant of one's own ignorance. In J. Olson & M. P. Zanna (Eds.), *Advances in experimental social psychology* (Vol. 44, pp. 247–296). New York, NY: Elsevier.

Dunning, D. (2014a). *Can ideas be too "smart" for their own good?* Unpublished manuscript. Cornell University, Ithaca, NY.

Dunning, D. (2014b). *The irony of advice: Only knowledgeable recipients can separate good from bad advice*. Unpublished manuscript, Cornell University, Ithaca, New York.

Dunning, D. (2014c). The problem of recognizing one's own incompetence: Implications for self-assessment and development in the workplace. In S. Highhouse, R. S. Dalal, & E. Salas (Eds.), *Judgment and decision making at work* (pp. 37–56). New York, NY: Taylor & Francis.

Dunning, D. (2014d). *Social comparison goes awry: How lopsided accuracy in social judgment underlies overly flattering self-assessments*. Unpublished manuscript. Cornell University, Ithaca, NY.

Dunning, D., & Cone, J. (2014). *Does genius hide in plain sight?* Unpublished manuscript. Cornell University, Ithaca, NY.

Dunning, D., & Covington, C. (2014). *On the mediocrity of crowds: Judgment aggregation dismisses smart but unpopular ideas*. Unpublished manuscript. Cornell University, Ithaca, NY.

Dunning, D., Heath, C., & Suls, J. (2004). Flawed self-assessment: Implications for health, education, and the workplace. *Psychological Science in the Public Interest*, *5*, 69–106.

Dunning, D., Johnson, K., Ehrlinger, J., & Kruger, J. (2003). Why people fail to recognize their own incompetence. *Current Directions in Psychological Science*, *12*, 83–86.

~~Dunning, D., & Roh, S. (2014). *False belief prevents insight into deficits in expertise*. Unpublished manuscript. Cornell University, Ithaca, NY.~~

~~Edwards, R. K., Kellner, K. R., Sistrom, C. L., & Magyari, E. J. (2003). Medical student self-assessment of performance on an obstetrics and gynecology clerkship. *American Journal of Obstetrics and Gynecology*, *188*, 1078–1082.~~

~~Ehrlinger, J., & Dunning, D. (2003). How chronic self-views influence (and potentially mislead) assessments of performance. *Journal of Personality and Social Psychology*, *84*, 5–17.~~

Ehrlinger, J., Johnson, K., Banner, M., Dunning, D., & Kruger, J. (2008). Why the unskilled are unaware? Further explorations of (lack of) self-insight among the incompetent. *Organizational Behavior and Human Decision Processes*, *105*, 98–121.

Ferraro, P. J. (2006). *Know thyself: Competence and self-awareness*. Unpublished manuscript. Georgia Tech University, Atlanta, GA.

Ferraro, P. J. (2010). Know thyself: Competence and self-awareness. *Atlantic Economic Journal*, *38*, 183–196.

Ferraro, P. J., & Taylor, L. O. (2005). Do economists recognize an opportunity cost when they see one? A dismal performance from the dismal science. *B.E. Journal of Economic Analysis and Policy*, *4*, Article 7. AU:6

Findlay, R., & Kierzkowski, H. (1983). International trade and human capital: A simple general equilibrium model. *Journal of Political Economy*, *91*, 957–978.

Freund, P. A., & Kasten, N. (2012). How smart do you think you are? A meta-analysis on the validity of self-estimates of cognitive ability. *Psychological Review*, *138*, 296–321.

Friedrich, J. (1996). On seeing oneself as less self-serving than others: The ultimate self-serving bias? *Teaching of Psychology*, *23*, 107–109.

Galton, F. (1907, March 28). The ballot-box. *Nature*, *75*, 509–510.

Gardner, P. H., & Berry, D. C. (1995). The effect of different forms of advice on the control of a simulated complex system. *Applied Cognitive Psychology*, *9*, S55–S79.

Hacker, D. J., Bol, L., Horgan, D. D., & Rakow, E. A. (2000). Test prediction and performance in a classroom context. *Journal of Educational Psychology*, *92*, 160–170.

Hansford, B. C., & Hattie, J. A. (1982). The relationship between self and achievement/ performance measures. *Review of Educational Research*, *52*, 123–142.

Harries, C., Yaniv, I., & Harvey, N. (2004). Combining advice: The weight of a dissenting opinion in the consensus. *Journal of Behavioral Decision Making*, *17*, 333–348.

Harris, M. M., & Schaubroeck, J. (1988). A meta-analysis of self-supervisor, self-peer, and peer-supervisor ratings. *Personnel Psychology*, *41*, 43–62.

Haun, D. E., Zeringue, A., Leach, A., & Foley, A. (2000). Assessing the competence of specimen-processing personnel. *Laboratory Medicine*, *31*, 633–637.

Helzer, E. G., & Dunning, D. (2012). Why and when peer prediction is superior to self-prediction: The weight given to future aspiration versus past achievement. *Journal of Personality and Social Psychology*, *103*, 38–53.

Hodges, B., Regehr, G., & Martin, D. (2001). Difficulties in recognizing one's own incompetence: Novice physicians who are unskilled and unaware of it. *Academic Medicine*, *76*, S87–S89.

Krajč, M., & Ortmann, A. (2008). Are the unskilled really that unaware? An alternative explanation. *Journal of Economic Psychology*, *29*, 724–738.

Krueger, J., & Mueller, R. A. (2002). Unskilled, unaware, or both? The contribution of social-perceptual skills and statistical regression to self-enhancement biases. *Journal of Personality and Social Psychology*, *82*, 180–188.

Kruger, J., & Dunning, D. (2002). Unskilled and unaware – But why? A reply to Krueger and Mueller. *Journal of Personality and Social Psychology*, *82*, 189–192.

Kruger, J. M., & Dunning, D. (1999). 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*, 1121–1134.

List C., & Goodin, R. (2001). Epistemic democracy: Generalizing the Condorcet Jury Theorem. *Journal of Political Philosophy*, *9*, 277–306.

Littlepage, G. E., Schmidt, G. W., Whisler, E. W., & Frost, A. G. (1995). An input-process-output analysis of influence and performance in problem-solving groups. *Journal of Personality and Social Psychology*, *69*, 877–889.

Mabe, P. A., III, & West, S. G. (1982). Validity of self-evaluation of ability: A review and meta-analysis. *Journal of Applied Psychology*, *67*, 280–286.

Malkiel, B. G. (2012). *A random walk down wall street: The time-tested strategy for successful investing* (10th ed.). New York, NY: Norton & Co.

Mannes, A. E., Soll, J. B., & Larrick, R. P. (2014). The wisdom of select crowds. *Journal of Personality and Social Psychology*, *107*, 276–299.

Marshall, A. (1890). *Principles of economics*. London: Macmillan.

Mellers, B., Ungar, L., Baron, J., … Tetlock, P. (2014). Psychological strategies for winning a geopolitical forecasting tournament. *Psychological Science*, *25*(5), 1106–1115.

Metcalfe, J., & Shimamura, A. P. (1994). *Metacognition: Knowing about knowing*. Cambridge, MA: MIT Press.

~~Miller, T. M., & Geraci, L. (2011). Training metacognition in the classroom: The influence of incentives and feedback on exam predictions. *Metacognition & Learning*, *6*, 303–314.~~

Mincer, J. (1974). *Schooling, experience, and earnings*. New York, NY: Columbia University Press.

Miner, F. C. (1984). Group versus individual decision making: An investigation of performance measures, decision strategies, and process losses/gains. *Organizational Behavior and Human Performance*, *33*, 112–124.

Moore, A. (2013). Google admits its famous job interview questions were a 'complete waste of time'. *Death and Taxes Magazine*, June 20. Retrieved from http://www.deathandtaxesmag.com/200732/google-admits-its-famous-job-interview-questions-were-a-complete-waste-of-time/

Morris, E. (2010). The anosognosic's dilemma: Something's wrong but you'll never know what it is (Part 1). *Business Insider*, June 20. Retrieved from http://opinionator.blogs.nytimes.com/2010/06/20/the-anosognosics-dilemma-1/?_php=true&_type=blogs&_r=0

Moss, C. (2014). *17 former Google interview questions so ridiculous the company banned them*, July 8. Retrieved from http://www.businessinsider.com/hard-google-interview-questions-2014-7?op=1

Nagel, M. (2010). A mathematical model of democratic elections. *Current Research Journal of Social Sciences*, *2*, 255–261.

Nickerson, R. S., Baddeley, A., & Freeman, B. (1987). Are people's estimates of what other people know influenced by what they themselves know? *Acta Psychologica*, *64*, 245–259.

Park, Y.-J., & Santos-Pinto, L. (2010). Overconfidence in tournaments: Evidence from the field. *Theory and Decision*, *69*, 143–166.

Pronin, E., Lin, D. Y., & Ross, L. (2002). The bias blind spot: Perceptions of bias in self versus others. *Personality and Social Psychology Bulletin*, *28*, 369–381.

Schlösser, T., Dunning, D., Johnson, K., & Kruger, J. (2013). How unaware are the unskilled? Empirical tests of the "signal extraction" counter explanation for the Dunning-Kruger effect in self-evaluations of performance. *Journal of Economic Psychology*, *39*, 85–100.

Schultz, T. W. (1963). *The economic value of education*. New York, NY: Columbia University Press.

Sheldon, S., Dunning, D., & Ames, D. R. (2014). Emotionally unskilled, unaware, and uninterested in learning more: Biased self-assessments of emotional intelligence. *Journal of Applied Psychology*, *99*, 125–137.

Sherif, M., & Hovland, C. I. (1961). *Social judgment: Assimilation and contrast effects in communication and attitude change*. New Haven, CT: Yale University Press.

Simons, D. J. (2013). Unskilled and optimistic: Overconfident predictions despite calibrated knowledge of relative skill. *Psychonomic Bulletin & Review*, *20*, 601–607.

Surowiecki, J. (2004). *The wisdom of crowds: Why the many are smarter than the few and how collective wisdom shapes business, economies, societies, and nations*. London: Little, Brown.

Veloski, J. J., Fields, S. K., Boex, J. R., & Blank, L. L. (2005). Measuring professionalism: A review of studies with instruments reported in the literature between 1982 and 2002. *Academic Medicine*, *80*, 366–370.

Williams, E. F., Dunning, D., & Kruger, J. (2013). The hobgoblin of consistency: Algorithmic judgment strategies underlie inflated self-assessments of performance. *Journal of Personality and Social Psychology*, *104*, 976–994.

Yaniv, I., & Kleinberger, E. (2000). Advice taking in decision making: Egocentric discounting and reputation formation. *Organizational Behavior and Human Decision Processes*, *83*, 260–281.

Zell, E., & Krizan, Z. (2014). Do people have insight into their abilities? A metasynthesis. *Perspectives on Psychological Science*, *9*, 111–128.

Zenger, T. R. (1992). Why do employers only reward extreme performance? Examining the relationships among performance, pay, and turnover. *Administrative Science Quarterly*, *37*, 198–219.

UNCITED REFERENCES

AU:5

Caputo and Dunning (2005); Dunning and Roh (2014); Edwards, Kellner, Sistrom, and Magyari (2003), Ehrlinger and Dunning (2003), Haun, Zeringue, Leach, and Foley (2000), Miller and Geraci (2011), Park and Santos-Pinto (2010), Williams, Dunning, and Kruger (2013)

AUTHOR QUERY FORM

	Book: AGP-V032-3611074 **Chapter: CH006**	**Please e-mail or fax your responses and any corrections to:** **E-mail:** **Fax:**

Dear Author,

During the preparation of your manuscript for typesetting, some questions may have arisen. These are listed below. Please check your typeset proof carefully and mark any corrections in the margin of the proof or compile them as a separate list.

Disk use

Sometimes we are unable to process the electronic file of your article and/or artwork. If this is the case, we have proceeded by:

☐ Scanning (parts of) your article ☐ Rekeying (parts of) your article

☐ Scanning the artwork

Bibliography

If discrepancies were noted between the literature list and the text references, the following may apply:

☐ **The references listed below were noted in the text but appear to be missing from your literature list. Please complete the list or remove the references from the text.**

☐ ***UNCITED REFERENCES:* This section comprises references that occur in the reference list but not in the body of the text. Please position each reference in the text or delete it. Any reference not dealt with will be retained in this section.**

Queries and/or remarks

Location in Article	Query / remark	Response
AU:1	(Dunning & Kruger, 1999) is modified as (Kruger & Dunning, 1999) in the abstract, as appearing in the reference list. Please confirm.	
AU:2	The following references are cited but not present in the reference list: "Kay & Shipman, 2014; Toplack & Stanovich, 2002." Please provide complete reference details.	

AU:3	Please check whether the year in (Dunning, 2014) refers to 2014a, 2014b, 2014c, or 2014d as appearing in the reference list.	
AU:4	In the artwork of Fig. 4, please spell out “Pct.” for clarity.	
AU:5	The following references are not cited in the text: “Caputo and Dunning (2005); Dunning and Roh (2014); Edwards, Kellner, Sistrom, and Magyari (2003); Ehrlinger and Dunning (2003); Haun, Zeringue, Leach, and Foley (2000); Miller and Geraci (2011); Park and Santos-Pinto (2010); Williams, Dunning, and Kruger (2013).” Please cite them and if no citation is available then these references need to be removed from the list.	
AU:6	Please check whether reference Ferraro & Taylor (2005) is okay set. Also, please provide page range.	