Blinded by “Insight”

Self-Assessment and Its Role in Performance Improvement

Kevin W. Eva, Glenn Regehr, and Larry D. Gruppen

Beware the Jabberwock, my son!
The jaws that bite, the claws that catch!
Beware the Jabberwock, my son!
The frumious Bandersnatch!

LEWIS CARROLL, Through the Looking Glass

In Carroll’s Through the Looking Glass Alice discovers a poem, “Jabberwocky,” that can be read only when held up to a mirror. So too have many others suggested that we as individuals can read ourselves only through deliberate acts of reflection. Among professionals ranging from athletic coaches to business leaders, there is a general belief that the path to better performance is through “looking in the mirror” to openly and honestly identify one’s weaknesses and take steps to improve them. As a result, many models of professional self-regulation worldwide and many health professional training curricula have incorporated some form of planned self-assessment activity into their practices (Sargeant et al. 2010). In fact, the health professions may provide the most extreme example, as their current models of maintenance of competence and self-regulation are
formalizations, in large part, of the instruction "Physician, know thyself." Generally, the practitioner is expected to reflect on his or her performance (preferably by incorporating indicative evidence) and to determine an appropriate course of action to ensure self-directed learning and maintenance of competence (Handfield-Jones et al. 2002). This expectation has prompted an industry of research into self-assessment and related constructs that has yielded both data and a reaction reminiscent of Alice's after she has read the poem. "It seems very pretty," she said when she had finished it, "but it's rather hard to understand!" (You see she didn't like to confess, even to herself, that she couldn't make it out at all.) "Somehow it seems to fill my head with ideas—only I don't exactly know what they are!!" 

Debate regarding the extent to which self-assessment can or should play a central role in the regulatory efforts and competency frameworks of the health professions has amplified in recent years (Eva and Regehr 2005). Most of it has centered around whether individuals are inherently incapable of judging the strength of their abilities or whether such judgments tend to correlate poorly with externally derived measures because individuals need to learn how to be reflective, how to incorporate external feedback into their self-perceptions, and what the community's standards are with respect to performance expectations. Debate at this level implies that the fundamental challenge in the notion of self-regulation resides in the ability of individuals to perceive performance deficiencies accurately. Unquestionably we each have more information available to us with which to judge our own abilities than do any external observers. It is this wealth of information, however, that may prevent us from generating accurate impressions of our own abilities. Our research and that of others raises questions about the adequacy of self-assessment, reasons why inadequate self-assessment may actually be adaptive, and the many ways in which we can fool ourselves into believing that we have privileged insight into our own capacities. This research raises fundamental concerns about the way in which the health professions typically conceive of self-assessment and the purposes for which it can be applied in a productive manner. Further, it is becoming increasingly clear that even if individuals were readily able to identify strengths and weaknesses in their performance, there are other issues that must be considered to determine the extent to which individuals can be expected to use perceived deficits to guide performance improvements (Regehr and Eva 2006; Regehr and Mylopoulos 2008). 

In this case, the self-improvement imperative is how health care organizations conclude the regulatory process and accumulate the relevant data. 

Laying the Foundation 

Judgments about whether self-assessment can or should play a central role in the regulatory efforts and competency frameworks of the health professions has amplified in recent years (Eva and Regehr 2005). Most of it has centered around whether individuals are inherently incapable of judging the strength of their abilities or whether such judgments tend to correlate poorly with externally derived measures because individuals need to learn how to be reflective, how to incorporate external feedback into their self-perceptions, and what the community's standards are with respect to performance expectations. Debate at this level implies that the fundamental challenge in the notion of self-regulation resides in the ability of individuals to perceive performance deficiencies accurately. Unquestionably we each have more information available to us with which to judge our own abilities than do any external observers. It is this wealth of information, however, that may prevent us from generating accurate impressions of our own abilities. Our research and that of others raises questions about the adequacy of self-assessment, reasons why inadequate self-assessment may actually be adaptive, and the many ways in which we can fool ourselves into believing that we have privileged insight into our own capacities. This research raises fundamental concerns about the way in which the health professions typically conceive of self-assessment and the purposes for which it can be applied in a productive manner. Further, it is becoming increasingly clear that even if individuals were readily able to identify strengths and weaknesses in their performance, there are other issues that must be considered to determine the extent to which individuals can be expected to use perceived deficits to guide performance improvements (Regehr and Eva 2006; Regehr and Mylopoulos 2008). 

In this case, the self-improvement imperative is how health care organizations conclude the regulatory process and accumulate the relevant data. 

Laying the Foundation 

Judgments about whether self-assessment can or should play a central role in the regulatory efforts and competency frameworks of the health professions has amplified in recent years (Eva and Regehr 2005). Most of it has centered around whether individuals are inherently incapable of judging the strength of their abilities or whether such judgments tend to correlate poorly with externally derived measures because individuals need to learn how to be reflective, how to incorporate external feedback into their self-perceptions, and what the community's standards are with respect to performance expectations. Debate at this level implies that the fundamental challenge in the notion of self-regulation resides in the ability of individuals to perceive performance deficiencies accurately. Unquestionably we each have more information available to us with which to judge our own abilities than do any external observers. It is this wealth of information, however, that may prevent us from generating accurate impressions of our own abilities. Our research and that of others raises questions about the adequacy of self-assessment, reasons why inadequate self-assessment may actually be adaptive, and the many ways in which we can fool ourselves into believing that we have privileged insight into our own capacities. This research raises fundamental concerns about the way in which the health professions typically conceive of self-assessment and the purposes for which it can be applied in a productive manner. Further, it is becoming increasingly clear that even if individuals were readily able to identify strengths and weaknesses in their performance, there are other issues that must be considered to determine the extent to which individuals can be expected to use perceived deficits to guide performance improvements (Regehr and Eva 2006; Regehr and Mylopoulos 2008).
In this chapter we examine the foundations of many modern theories of self-improvement, lay and otherwise, raising fundamental concerns about how health professions typically conceive of and apply self-assessment. We conclude the chapter with a description of how models of professional self-regulation can be effectively modified given the evidence base that has accumulated to date.

Laying the Groundwork

If you want to make the world a better place, take a look at yourself and then make a change.

MICHAEL JACKSON, "Man in the Mirror"

Judgments of personal competence underlie everything we do, and the notion that we can improve ourselves (and the world) by more carefully considering the "person in the mirror" is prevalent. We make judgments about our competence whenever we decide that we are able to navigate the streets in a strange city rather than rely on the support of professional taxi drivers, whenever we apply for a promotion to a job that we consider ourselves capable of doing, or whenever we initiate a treatment regimen for patients we consider ourselves knowledgeable or skilled enough to help. Through these examples it can be seen that the stakes of inaccurate self-assessment vary and that the decision not to act when one has the capacity can be as costly as the decision to act when one should not. Within the health professions it has largely become universal that the onus of responsibility for the maintenance of competence is placed on the individual practitioner. In scanning the revalidation and recertification policies of eighteen countries, Peck et al. (2000, 435), concluded that "approaches differ widely around the world, but most rely on professional self-regulation." As such, it is imperative that we understand the limits of the various activities related to maintenance of competence and professional development, including our ability to self-assess.

It would seem self-evident that each of us, as individuals, has greater opportunity to generate insights into the strengths and limits of our own performance and ability than does anyone else. There must be tremendous insight to be gained from total knowledge of one's past experiences (as opposed to the samples available to external observers), awareness of the
motivations leading us to act in certain ways, and the inner thoughts that others have access to only to the extent that we choose to (or can) share them. This notion is so self-evident and the intuition so powerful that we readily and commonly attribute flaws in others to lack of self-awareness; we easily discount information received from other sources as being noncredible (e.g., resulting from inadequate observation, personality conflicts, or invalid testing protocols); and we routinely present improved self-reflection as a mechanism to guide performance improvements. Any one of these interpretations will be correct in some circumstances, but the important issue is the proportion of times in which they are correct. As we have written before, accurately knowing that one could never play professional football is not an argument for accurate self-assessment because one needs also to consider how frequently we are inaccurate in “knowing” we are able to (or could not) perform certain tasks well (Eva and Regehr 2008).

Before getting into these issues in greater detail, however, it is crucially important to clearly and explicitly define the terms we will use to consider the issues we hope to address because we are convinced that the implicit nature of these definitions in many conversations has yielded considerable confusion and conflict. We make no claim that these are the only ways in which the following terms could be used, but they are definitions that have helped us conceptualize critical distinctions between aspects of cognition and models of self-regulation that are often confounded in the health professions education literature. Furthermore, we think it critical that the community come to some form of agreed-on taxonomy to lessen the communication difficulties, misunderstandings, and derivation of inappropriate implications that can arise from loose use of language.

Consistent with our earlier work, we will define self-assessment as a personal, unguided reflection on performance for the purposes of generating an individually derived summary judgment of one’s own level of knowledge, skill, and understanding in a particular area (Eva and Regehr 2008). There are several aspects of this definition that we consider important to highlight. First, in this definition we have framed self-assessment as a process, not as the outcome of that process. We think this distinction is important because the failure to make this distinction has been one of the major sources of confusion in the community’s discussions. Several authors have suggested, for example, that manipulations that bring self-assessment scores in line with objectively derived scores are an indicator that they have

“improved” on their self-assessments (or vice versa). We contend that alignments between self-assessment and outcome measures should be seen as alignment in performance, not necessarily in knowledge or skill, where scoring may be complex (see Chapter 5).

Thus, while we agree with those who label this as “self-assessment,” we disagree with those who label those outcomes as “self-assessment” because they do not align with self-assessments in the way issue here because the terms in which they are described differ.

Second, self-assessment, as defined in this chapter, is defined as “a personal, unguided reflection on performance for the purposes of generating an individually derived summary judgment of one’s own level of knowledge, skill, and understanding in a particular area.” We believe it is more important to focus on the process of self-assessment, rather than its function. Self-assessment is a process that aligns with one’s self-assessment in the way it is described in this chapter. We think it is important to better personalize and differentiate self-assessment. Outside of this chapter, we believe that personalized self-assessment is more important, whereas in this chapter, we think it is more important.
the inner thoughts that we agree to (or can) share them. 

As powerful that we readily 

of self-awareness; we 

selves as being noncredible 

ility conflicts, or invalid 

ved self-reflection as a 

ny one of these inter- 

nt but the important issue 

et. As we have written 

ay professional football 

cause one needs also to 

wing” we are able to (or 

ehr 2008). 

, however, it is crucially 

we will use to consider 

ounced that the implicit 

as yielded considerable 

ese are the only ways 

are definitions that 

etween aspects of cogni-

confounded in the health 

ink it critical that the 

onomy to lessen the con-

ervation of inappropri-

age 

self-assessment as a per-

purposes of generating 

one’s own level of knowl-

(Eva and Regehr 2008). 

we consider important 

self-assessment as a 

think this distinction is 

ication has been one of the 

ussions. Several authors 

bring self-assessment 

indicator that they have 

“improved self-assessment.” Yet this interpretation is problematic. Since 

most people tend to overassess their performance, improving actual 

formance tends to bring actual scores more in line with self-assessment 

scores (Kruger and Dunning 1999, study 4). However, this incidental 

ignment of self-assessed and actual scores does not imply an improve-

ient in the process of self-assessment any more than moving a target to 

where someone is throwing a ball implies that we are improving their aim. 

Thus, we feel it important to distinguish self-assessment the verb from self-

essment the noun. Unfortunately, we have as yet found no sensible set of 

ables that readily enables this distinction; so we will simply highlight the 

issue here and, throughout this chapter, try to be explicit about whether we 

are describing the process of self-assessment or the outcome of that process. 

Second, self-assessment as we have defined it is distinct from issues of 

self-concept (Cervone 2000) in that self-assessment involves an element 

of context-specific appraisal rather than simply indicating how one self-

identifies. Relative to self-concept, self-assessment, in the way we define it, 

is more immediate and transient rather than a longer-term stable construc-

tion of our self-worth. As a result, the construct of self-assessment better 

aligns with self-efficacy, the domain-specific belief in one’s capabilities to 

perform in a certain manner (Cervone 2000). In the social psychology 

literature, self-efficacy is spoken of as a determinant of performance in addi-

tion to being a judgment of one’s ability (i.e., positive self-efficacy can yield 

better performance relative to negative self-efficacy; see Dweck 1999). 

Outside of that literature, however, self-assessment is more generally spoken 

of simply in terms of reflecting one’s capability and, as a result, while we 

think the bidirectional aspect of self-appraisal and performance is im-

portant, we will treat the terms self-assessment and self-efficacy separately. 

Third, the “individually derived” portion of our definition of self-assessment 

stands in stark contrast to the definitions offered by Boud (1999) and 

by Epstein, Siegel, and Silberman (2008), who argued that self-assessment 

should be conceived of as self-appraisal derived from the interaction be-

between internally derived perceptions and externally generated data regard-

ing the strength of one’s performance or abilities. We do not deny that such 

interactions are important and, indeed, we presume that self-assessments 

are always informed to some extent by one’s experienced successes and 

failures. However, we have included the concept “individually derived” 

in our definition for two reasons. First, much of the research addressing
self-assessment has implicitly used this concept in the operational definition. That is, most studies attempt to measure self-assessment ability by having participants engage in some performance, assessing that performance, asking participants to make their own summary assessment of the performance, then comparing the externally derived assessments to the individually derived assessments. We believe that the ubiquity of this "guess your grade" model of self-assessment research (and its implied assumptions about the process) both reflects and reinforces an implicit model of the self-assessment process that leans heavily on a belief that people can determine for themselves their level of performance and ability. Second, we believe that important cognitive processing takes place at the intersection between one's self-perceptions and the interpretation of external data (as will be discussed below). This processing is heavily determined by factors beyond the data on which the self-assessment process is purportedly based. Thus, we will use the term informed self-assessment (Sargeant et al. 2010) to refer to Boud's and Epstein, Siegel, and Silberman's versions of self-assessment as a deliberately interactive process of incorporating data available from the environment into one's self-assessment. In this way, informed self-assessment is related to self-directed assessment seeking (Eva and Regehr 2008), the pedagogical activity of intentionally looking outward for formative and summative assessments of one's current levels of performance.

Finally, we note that the definition of self-assessment as resulting in a "summary judgment" of one's ability in a particular area invokes a crucial distinction between broader judgments of how well one performs in an area generally and how one determines, in the moment, whether or not one is performing in a way that is yielding the outcomes one desires (what we have referred to in previous work as self-monitoring: Eva and Regehr 2007, 2011). That is, as a broad judgment, self-assessment might be used as a mechanism through which to self-limit one's decisions about general domains of practice (e.g., "I no longer accept referrals for those types of cases"). As a much more narrow level of judgment, self-monitoring can be used, in the moment, to determine whether one needs help with a particular task (e.g., "I should consult my colleague about the implications of that test result"). In this way self-assessment and self-monitoring align well with the reflective discourse identified by Hodges in chapter 1 and Schön's (1983) notions of reflection-on-practice and reflection-in-practice, respectively. In a broader sense of the literature, we may also understand self-assessment as a reflection of self.

Direct Teaching...

In 1962 the Physician Training Medical Commission of Canada, which each year inquires into the matter of training conducted, reviewed the conclusions that people do not rate their own work as the topic with the lowest latitude up, and only be low to the extent that professionals are not related to the work of external factors. This damage, in our language, becomes a problem (the second dimension of medical education).

Because, in 2006, for a variety of reasons, we see a pattern of training and team reporting that performance indicators and evidences and...
respectively. We have opted not to adopt these terms because reflection is a broader notion in that it has also been used (in much of the expertise literature, for example) to describe the act of investing mental energy into understanding “why” the world works in the way it does rather than simply as a mechanism through which to judge personal competence (Mann, Gordon, and MacLeod 2009).

Direct Tests of Self-Assessment

You’re on your own. And you know what you know. And YOU are the guy who will decide where to go.

Dr. Seuss, Oh, the Places You’ll Go!

In 1962 the American Medical Association and Association of American Medical Colleges issued a joint report that declared, “Self-appraisal is a test which each of us has applied to himself for so long that we hardly consider the matter” (Gordon 1991, 762). Since then hundreds of studies have been conducted to “consider the matter,” and their results have generally led to the conclusion that, perhaps for the only time, Dr. Seuss was wrong: people do not seem to “know what they know.” In one of the first reviews on the topic within the health professions, Gordon (ibid.) summarized the literature up to 1990: “The validity of self-assessed performance was found to be low to moderate and did not improve with time in conventional health professions training programs. Self-assessed performance seemed closely related to generalized self-attributions and was minimally influenced by external feedback in the form of test scores, grades, or faculty assessments.” This damning conclusion did not curb the optimism Gordon expressed regarding the potential to create effective interventions to overcome this problem (1992), but despite two decades of additional research, the conclusions of many others are consistently bleak.

Because the literature has been reviewed many times (see Davis et al. 2006, for an example focused on self-assessment among practicing physicians), we will not provide a comprehensive review here. That said, the pattern of findings can be summarized fairly concisely: an investigator team reports a poor correlation between self-assessment and external performance data; subsequent investigators identify methodological weaknesses and run a new study prior to reporting the same poor correlations;
repeat. To be clear, this representation of the literature is not a criticism of others as we each spent considerable time and effort on this treadmill (Eva et al. 2004; Ward, Gruppen, and Regehr 2001). The phenomenon of a poor correlation between self-rated performance and other forms of assessment is as robust as the intuition that we are able to self-assess, and therefore as a community we seem to have concluded that the data must be wrong. Yet inaccurate self-assessment has been shown across many domains, both within and outside the health professions (Dunning, Heath, and Suls 2004). There is variability in the strength of the correlations observed (Ward, Gruppen, and Regehr 2001), but no more than one would expect of any normal distribution. There is some indication that correlations can be improved with intensive feedback (yielding Gordon's optimism), but these generally appear more indicative of individuals being able to learn about weaknesses in a particular performance rather than improved capacity for accurate self-assessment (Eva 2001). More fundamentally, it is fair to question whether such correlational studies adequately represent the construct of self-assessment as opposed to a capacity to “guess your grade” (Collier, Verhulst, and Barrows 2005). In fact, it is studies that deviate from this norm that have perhaps been most informative in yielding understanding of the difficulties inherent in judging oneself.

In a landmark study that won an Ig Nobel Prize in 2000 for “research that first makes people laugh and then makes people think,” Kruger and Dunning (1999) demonstrated that those who are unskilled in a particular task (e.g., grammar and composition) have what amounts to a perceptual blindness to the aspects of the task that define skilled performance. Participants of all skill levels generally rated themselves as above average, thereby reproducing the common phenomenon that we tend to believe we perform better than most in socially desirable domains (commonly referred to as the Lake Wobegon effect), underestimate how long it will take us to complete tasks, and are overconfident in our judgments and predictions (Dunning, Heath, and Suls 2004). When participants were offered feedback via exposure to the test responses of a handful of their colleagues, those who performed particularly well on the grammar test (and tended to somewhat underestimate their ability) were able to recognize the divergence between their perceptions and reality as evidenced by their raising the ratings they gave their performance.
More important, however, is that those who performed particularly poorly (and tended to considerably overestimate their ability) were unable to take advantage of this additional information and did not alter their self-ratings. In other words, even though their colleagues demonstrated better grammar, poor performers did not recognize this fact. A later study presented in the same article demonstrated that intervening to improve performance (i.e., giving participants the skill) did improve the alignment between self-assessment and actual performance because the latter rose to meet the former. The main findings from this work have been replicated in a clinical setting focused on communication skills performance during an objective structured clinical examination (OSCE) (Hodges, Regehr, and Martin 2001).

Contrary to much of the rhetoric within the health professions, these findings suggest that the path to improved performance is directed training and feedback, not better self-assessment. Further, they also force us to recognize that alignment between self-ratings and actual performance does not necessarily derive from capacity to self-assess. Finally, the results remind us of the ever-present phenomenon of context specificity in that the extent to which one can self-assess well appears to be strongly context bound (Eva 2003; Fitzgerald, White, and Gruppen 2003). We assess ourselves accurately in domains in which we have skill/knowledge, but not in domains in which we do not. Unfortunately (or perhaps fortunately, for reasons that will be presented later), the implication is that we have no way of knowing for ourselves when we are in one state or the other.

Enough evidence has accumulated in both laboratory settings and the real world that is supportive of the argument that one cannot rely exclusively on self-perceptions of performance or ability that we will not belabor that point any further (Dunning, Heath, and Suls 2004). Instead, we turn our attention here to research aimed at fleshing out what to do with this information. Before doing so, however, it is worth explicitly noting that we have deliberately selected the pronouns "we" and "us" throughout this chapter when discussing difficulties with self-assessment. Because we are fans of irony, one of our favorite results from this entire domain of study is that people tend to believe they are less prone to bias in their self-assessment than they believe their peers to be (Friedrich 1996; Pronin, Lin, and Ross 2002). Poor self-assessment is not just a medical student problem, it is not just a health professional problem, and it is certainly not just
someone else’s problem; it impacts all of us in every aspect of our daily activities. As a result, determining what to do about the findings that have accumulated to date in this domain is not straightforward. The very difficulties inherent in internalizing the counterintuitive position that our self-assessments are flawed can cause us to be blinded by the insights they provide, thereby making it difficult to fully use the external information made available to us.

The Psychological Mechanisms Influencing Self-Assessment and Our Belief in It

It would seem to me
I remember every single…thing I know.

The Tragically Hip, “At the Hundredth Meridian”

We believe the reason so much time and energy has been devoted to grappling with issues of self-assessment and the role it should play in professional self-regulation is that it is so highly counterintuitive that one would struggle to self-assess accurately and so seemingly easy to attribute difficulties with others to their lack of this capacity. As a species, we may not readily admit our flaws to others, but surely we are aware of them at some fundamental level. And if we are not, does that not shake the very core of who we are as humans? Thoughts resonate in our heads, we can logically reason our way from one state to another, and we can plan and defend our actions. How could we be unaware of the limits of our ability when we are so seemingly capable of sophisticated introspection? Well, what if that belief in our ability to self-assess (and, more generally, to introspect—see Bargh and Chartrand 1999) was just a by-product (i.e., an epiphenomenon) of the way we’ve been wired to think about the world? We do not wish to become overly philosophical in raising these questions, but they speak to the very root of what the field is trying to accomplish when suggestions are made that improving self-assessment is a means toward the ends of better lifelong learning, maintenance of competence, and patient safety.

Close your eyes for a moment and think about yourself driving. If you are like most people, you have just thought of a reasonably relaxed experience with few of the details that determine the success of one’s driving in the real world. Was the weather bad? Were there distractions present at the side of the road? How did we have drawn this conclusion? In general people do not have the time or energy to seek out the detail that is necessary, and thus the easiest way to make an assessment in a scenario is to intuit (see Tversky and Kahneman 1973) and use the information we have to point lines. This allows us to make it more likely that we are acquiring our overall sense from those cues, meaning that the idea of self-assessing, or anyone else’s, seems to be an intuitive judgment made from these cues. From these and our counterintuitive belief in confidence, we may wonder why it is that few would claim that the basis is sheathed in the actual events with this in mind. Philosophers talk about the idea of confidence (Eva et al. 2003) in contrast to the social phenomenon, in order to predict and understand ourselves, our strangers, our colleagues, Donald and colleagues (1984), have available their alleged access.

But why would we be so content? Why rule people in their self-assessment if we want to take a more external perspective to “fix the person” rather than the context and teaching? We consider...
the side of the road? Was there something playing on the radio that may have drawn your attention from your mirrors? We could go on, but the general point is that when asked to think of ourselves performing a task, we tend to think of the gist rather than the details even though it is the detail that will determine our success (Vallone et al. 1990). Furthermore, the easier it is to think of a scenario, the more likely we are to perceive that scenario as representing a probable state of affairs (Tversky and Kahneman 1973). Even the worst drivers more often than not get from point A to point B without disastrous consequences. Every time they do it makes it more likely that they will both infer from their driving record and the overall sense of safety derived from their reconstructed memories of driving that they are capable drivers and have difficulty understanding how anyone else, without this accumulation of evidence, could rightly pass negative judgment on their skill level. In other words, the “insight” derived from these past experiences simultaneously blinds us to our ability level and our capacity to judge our ability level. A trial judge may gain great confidence, over her years on the bench, in her ability to recognize liars because of the many times she has had to make these distinctions, but given that few who stand trial ever admit that they really were lying, on what basis is she deciding that her earlier judgments were accurate? Consistent with this idea, in a series of focus groups, we have heard health professionals talk about how their amount of experience determines their feelings of confidence with little to no mention of the quality of those experiences (Eva et al. 2011). A large variety of studies from the medical education and the social psychology literatures have suggested that others are better able to predict any number of things about us than we are able to predict about ourselves, including the likelihood of passing examinations (Risucci, Tortolani, and Ward 1989) and the length of our romantic relationships (MacDonald and Ross 1999), further suggesting that the extra information we have available to us about ourselves creates an illusion that we have privileged access to knowledge of our ability level.

But why would such errors in judgment be maintained? As a general rule people do not like to make mistakes or to be ignorant, and they tend to want to take action to overcome such states. Indeed, this tendency to want to “fix the problem” has led many to propose the need to create curricula and teaching strategies aimed at improving the accuracy of self-assessment. We consider this response misguided. Gilbert and Wilson (2000) coined
the term "psychological immune system" as a way of describing the many
cognitive mechanisms that protect us from threatening information. As
with our biological immune system, the idea is that, within limits, we are
better off maintaining an optimistic outlook on our ability and a "rose colored" self-concept. This is well documented in the psychology and psy-
chiatry literature, where it is an accepted finding that depressed people
appear to have a more realistic perception of their importance, reputation,
locus of control, and abilities than those who are not depressed (Alloy and
Abramson 1979), a phenomenon well enough established to have acquired
the label "depressive realism" (Dobson and Franche 1989). As a result, we
have built-in mechanisms to counteract external hazards that make us vul-
nerable (see Nussbaum and Dweck 2008). When we encounter informa-
tion that contradicts our positive preconceptions about ourselves we have a
tendency to automatically and unconsciously discount that information by
blaming our car accident, for instance, on the rain, discounting the health
risks inherent in our behaviors as being overblown by bad science, or blam-
ing our poor performance on a headache or distractions created by our
personal circumstances (Eva et al. 2011; Gilovich 1991; Tavris and Aronson
2007). Again, these interpretations will sometimes be accurate, but they
become problematic when they are overly reflexive—though, ironically,
such reflexivity may be a good thing.

This tendency may be fortunate because the self-efficacy literature leads
to the conclusion that self-assessments are not simply reflections of perfor-
mance. Instead, they should be thought of as fundamental determinants of
performance as there is an important reciprocity between self-efficacy and
success. Such beliefs are based on a variety of experimental studies that
have demonstrated that randomly lowering self-efficacy can result in per-
formance deficits relative to participants who have their self-efficacy raised
by praise (Cervone 2000). Importantly, however, the type of praise matters.
For example, Carol Dweck's (1999) work has illustrated that praising stu-
dents for their effort yields greater long-term gain than praising students
for their intelligence; those praised for their intelligence will more readily
give up on a difficult task than the former group because they infer that
they have reached the limit of their ability rather than perceiving a need
to try harder. Implicit understandings of the value of positive self-efficacy
are present throughout the professional identity literature in which indi-
viduals speak of the importance of feeling autonomous and confident in

their own ability to achieve personal and professional goals. As the adage
points out, "knowledge is power." Mann et al. (1995) describe this group of
participants as having a "need for interaction" or a "need for meaning in
practice".

In all, then, the construction and maintenance of a professional identity create
a web of interconnected, overlapping, and interdependent perspectives. It is
not enough to view the professional identity in isolation. Instead, notions of
knowledge, skills, and performance are all closely linked to the way we
conceptualize and experience the world around us. As medical professionals
habitually do, we can begin to see how professional identity construction
becomes a way to "frame" the world, that is, a way the world becomes

Challenges in Identity Formation

The anatomy of the professional identity construction process, then, is a
mechanism that shapes professionals' beliefs about their role and perfor-

assessment and performance. Knowledge, skills, and even goals are
breath, and personal identity. Without a professional identity, individ-
uals would be unable to compute individuals' self-concept, personal identity, and goals that


describing the many ways in which this was true. As within limits, we are not aware of the ability and a "rose colored glasses" bias in psychology and psychiatry, that depressed people have a lower self-importance, reputation, and self-esteem (Alloy and Mehl 1989). As a result, we are more likely to blame ourselves than when we encounter information. It is important that we acknowledge the health implications of self-assessment and self-appraisal data. Tavris and Aronson (1978) showed that self-appraisal is not accurate, but they also showed—though, ironically, inaccurately—that self-appraisal can lead to self-enhancement and self-fulfilling prophecy.

Self-efficacy literature leads to the hypothesis that self-efficacy is a determinant of performance because self-efficacy and performance are positively correlated. Experimental studies that manipulate self-efficacy can result in changes in performance, so the self-efficacy raised is the type of praise that matters. Empirical data have indicated that praising students for their efforts rather than praising students for their performance will more readily enhance self-efficacy because they infer that they are capable of achieving a goal if they perceive a need for self-efficacy. This is a natural and difficult to overcome as breathing itself (and, indeed, what else would we use it for?). In fact we have found that, on balance, individuals' self-appraisals have a greater impact on their generation of learning goals than does the external feedback they receive (Eva et al. 2010). During their own abilities. Many theorists argue that taking on the identity of the profession is as fundamental an aspect of competence as are the knowledge and skills associated with the profession (Lave and Wenger 1991; Mann 2011; Monrouxe 2010). Similarly, we have recently noted in focus groups conducted with health professional trainees and practitioners that participants feel a sometimes paralyzing tension between knowing they need feedback to improve and fearing information that disconfirms their practice (Mann et al. 2011).

In summary, the empirical findings surrounding self-assessment and the context/content-specific nature of its accuracy suggest that we cannot create good self-assessors. Further, the data shedding light on the psychological mechanisms for poor self-assessment and the derived theoretical perspectives that highlight the value of confidence suggest that we should not even try. This perspective is fundamentally at odds with the prevalent notion that we need to improve self-assessment in order to enable learners and practitioners to improve performance. Rather, we would argue that we can have the greatest impact on self-assessment by ignoring it and focusing instead on ways in which we can enable individuals to develop the habits of seeking trustworthy feedback, acting on it without allowing it to become a personal threat, and enabling others to share such feedback in a way that facilitates its use.

Challenges Inherent in Trying to Inform One's Self-Assessment

It is impossible for anyone to begin to learn that which he thinks he already knows.

Epictetus, Discourses

The argument that improved self-assessment does not offer a mechanism through which one can appropriately limit one's practice or guide performance improvements is not an argument against the use of self-assessment. On the contrary, we think the latter argument to be a waste of breath, as the speed and automaticity of the self-assessment process is as natural and difficult to overcome as breathing itself (and, indeed, what else would we use in its stead?). In fact we have found that, on balance, individuals' self-appraisals have a greater impact on their generation of learning goals than does the external feedback they receive (Eva et al. 2010). During
an OSCE established for the dual purposes of assessing student performance and offering an educational opportunity by enabling students to be observed and receive direct and immediate feedback on their clinical performance, the investigators asked students to evaluate the quality of the feedback they received and to appraise their own skill. After the OSCE, participants were asked to generate a list of learning goals and learning activities they identified as important for their continuing professional development that resulted from their experience in particular OSCE stations. The self-appraisals, quality of feedback ratings, and observer assessments (among other variables) were then submitted to a regression analysis to determine which factors, personal or external, most influenced students’ generation of learning goals. The informed self-assessments that were generated by the students themselves were far and away the strongest determinant of learning goal generation, with the opinions of the raters regarding the strength of the students’ performance having little effect. These findings reinforce the notion that external feedback is never delivered in a vacuum and that, as such, it is crucial to understand the context of what the learner believes about his/her own performance if one hopes to motivate the learner to receive and use any feedback provided.

As a result, the challenge of enabling performance improvements becomes even more daunting in that calls to establish mechanisms for delivering data to the individual, while valuable, often fail to recognize the difficulty inherent in overriding the strongly formed opinions we all hold about our strengths and weaknesses. The medical education world and knowledge translation industry have expended considerable energy and resources to develop audit and feedback strategies that examine physician practice patterns and provide feedback reports that indicate how one’s practice aligns absolutely with formal professional guidelines or relative to a distribution of one’s colleagues (Foy et al. 2005). The results of such efforts have been highly variable, in part because they generally treat the delivery of feedback as a dispassionate exercise in which the recipient is a disinterested observer whose only emotional investment is a desire to get better. We do not doubt for a second that every health professional (save, perhaps, a few notable exceptions) has a genuine desire to perform well, but we do think it important to recognize that being told one can improve inevitably carries the implication (if not more) that one has been performing at a suboptimal level. Is it any wonder then that people are more likely to believe the feedback they are being offered when suggestions are made of how to improve?

Sargent (2006) suggested that the context in which feedback is given is clearly important. Participants who were provided with feedback and who were met with a randomised, constructive commentary were far more likely to respond to the feedback and improve their performance and to appreciate that there was much more to learning than simply being told that they were not performing up to the mark. As a result, feedback is always more likely to be effective when the more constructive and less confrontational approach is used.

Somen et al. (2010) also highlighted the act of giving feedback. For example, the time and place between the feedback giving and feedback taking are critical and can impact on how feedback is received and the implications on the physician’s future practice. The researchers noted the need to deliver good feedback in a non-confrontational environment, free of distractions, and to do so immediately after the event so that the feedback delivered is more likely to be incorporated into future practice. The time and place of feedback are significant in ensuring that feedback is used effectively and not simply as a point of reference. The feedback given needs to be relevant to the context in which the feedback is given and should be framed in a way that encourages reflection and action.
to believe data that suggest they are performing well relative to data that suggest they are not?

Sargeant et al. (2003), in a study of the uptake of feedback provided in the context of a physician-oriented multisource feedback system, found that participants' agreement with the scores assigned to them by their colleagues were positively correlated with the mean of those scores. That is, the higher we are rated, the more readily we believe the rating. Boehler et al. (2006), in a randomized controlled trial comparing different forms of feedback, illustrated that students gave higher teacher ratings when they received complimentary feedback, relative to when the feedback was constructively critical, despite finding greater performance gains in the latter group. When faced with the cognitive dissonance induced by the competing ideas that one is performing well and that one needs to improve, something has to give. It is much easier, given the psychological immune system, to poke holes in the external data provided (my tutor doesn't like me; my practice is different; that was just a fluke) than it is to decide that one's self-perceptions are flawed. Again, this argument is not put forward to say that such reasoning is always wrong, but is meant simply to draw attention to the notion that the more incongruent external data are with one's perceptions of reality, the more likely one is to simply ignore or actively discount the data.

Sometimes the tensions created by the cognitive dissonance inherent in the act of informing self-assessment are explicit and easily recognizable. For example, Kennedy et al. (2009) report that individuals feel tension between the desire to learn and concern about looking incompetent, thereby making people reluctant to seek feedback. Mann et al. (2011), in a study of physicians, medical students, and midwifery students, report similar findings and note further that individuals often struggle between wanting to deliver genuine and beneficial feedback while being hesitant to do so out of fear of damaging relationships. We suspect awareness of the emotionally evocative nature of feedback drives this tension. At other times, however, the tensions are felt and resolved much less consciously, further reinforcing the difficulty we have with knowing when we are right in discounting external information and when we are wrong to do so.

The social cognition literature presents the notion of “cold cognition,” the limitations of memory and attention that can lead us to reinterpret data for the sake of reconciling it with our established view of the world even in the absence of the shock or grief that may arise when negative data that
conflict with our worldview are presented (Eva et al. 2011). As examples, gamblers are known to interpret their losses as “near wins” (Gilovich 1991), and family physicians have been seen to attribute poor performance to extenuating circumstances like headaches while readily attributing good performances to their own ability (Eva et al. 2011). The universality of these sorts of discounting mechanisms and the resilience of our impressions indicate that we cannot simply tell people to be more open to external information or to collect data that will inform their self-assessments any more than we can successfully tell people to not be biased in their decision making. As regulatory authorities exist that are responsible for ensuring professional accountability, they can certainly mandate that people participate in some sort of reflective activity and members of the profession will go through the motions (Hays and Gay 2011). However, if we are to hope that the provision of data will influence performance, thereby fulfilling the formative purpose that self-assessment advocates suggest is paramount, then we need to determine how to create authentic activities that will inform individuals about how to improve without expecting to do so through an improved capacity to self-assess.

Where Do We Go from Here?

Now, in order to answer the question, “Where do we go from here?” which is our theme, we must first honestly recognize where we are now.

MARTIN LUTHER KING, “Where Do We Go from Here?”

It is not lost on us that to this point in the text we have likely raised problems more than we have provided solutions. This approach was felt necessary, however, because there are no simple solutions; because many of the solutions that can be proposed fall prey to the same cognitive constraints as the intuition that we can self-assess; and, as Dr. King said, because it is important to understand where we are now if we are to determine where we go from here. Given the theme of the discussion so far, it is presumably clear that we do not think we can “honestly recognize where we are now” when left to our own devices, but we have attempted to present samples of data from a variety of literatures to make three fundamental arguments:

1. The accuracy of self-assessment as an introspective process is too context dependent to be a trustworthy guide to performance improvement.
2. The inability to accurately self-assess our own weaknesses is adaptive in that the resulting optimism promotes tenacity and pushes personal and professional performance beyond the point that may be reasonable based on one's actual competence.

3. The same cognitive processes that impede the accuracy of self-assessment (and create a belief in our capacity to self-assess) prevent us from simply adopting external data when it conflicts with our self-assessment. This dynamic thereby limits the likely impact of simply informing self-assessment through the presentation of external data.

These three arguments lead us to the position that although we cannot trust self-assessment, we should encourage it as a way for peers and educators to discern how to deliver feedback in a manner that will allow receptivity and performance change. While that is not a guaranteed solution, we are led toward the view that feedback must be delivered from a position of understanding the recipients’ perspective to avoid it being so incongruent that it falls on deaf ears. It must also be delivered from a position of understanding that feedback inherently creates a threat to the individual. For these reasons, we argue below that the best we can do as educators and regulatory authorities is to create opportunities for individuals to discover ways in which their performance can be improved rather than divining strategies to directly alter self-assessment as a mechanism for performance improvement. In this final section we will elaborate on these points by constructing our best guess at how models of professional self-regulation and performance improvement can be effectively modified given the evidence base that has accumulated to date. This section, however, is necessarily more speculative than those that preceded it and, as a result, should be read more as a series of ideas and research findings we think to be important for moving us forward, rather than as a fully refined model.

While the notions of informed self-assessment and the need to separate self-assessment and performance improvement lead us to the literature on feedback interventions, that literature has been developing for over a century (Brand and Stratton 1905), making it impossible for us to provide a comprehensive review within the confines of this chapter. Fortunately, good reviews exist (Archer 2010; Hattie and Timperley 2007; Kluger and DeNisi 1996; Kluger and van Dijk 2010; Shute 2008) that have highlighted intriguing findings to help further our discussion of the integration of individuals’ self-assessment and external feedback. In a comprehensive meta-analysis, Kluger and DeNisi found that, contrary to the accepted wisdom
that feedback is a good thing, the more threatening feedback is to the recipient's self-concept, the more likely it is to have a negative impact on performance.

These findings are consistent with Dweck's (1999) data in that being able to attribute failure to insufficient effort would seem less of an affront to one's self-concept than being faced with the prospect that failure indicates the limits of one's intellectual capacity. Furthermore, the findings suggest a mechanism through which one can understand the relative impact of qualitative versus quantitative feedback provision that was illustrated by Butler (1987). In a randomized controlled trial, Butler presented feedback to individuals in one of three forms: quantitative (i.e., marks), qualitative (i.e., comments), or both together. Not surprisingly, given the literature illustrating the inadequacy of self-assessment, constructive criticism offered in the form of qualitative commentary yielded greater learning gains than did numerical feedback; if one cannot self-assess then one cannot determine how to do better based on a number alone. More interesting, however, was that participants given both quantitative and qualitative feedback subsequently improved only at the rate of those given quantitative feedback in isolation. While these findings have yet to be replicated in the health professions, they (along with other empirical results) led Kluger and DeNisi to conclude that feedback that draws attention to cues about where one stands in relation to others, thus threatening the self, is more likely to be detrimental than feedback that does not. That said, our personal experience suggests that students desire knowledge of their grade. How this tension can be resolved is yet to be determined.

One excellent example of how data might be used to enhance performance while effectively circumventing summative self-evaluations has been offered by Galbraith, Hawkins, and Holmboe (2008) in their description of the potential use of electronic patient records (EPRs) to improve practice. As they describe, if EPRs were structured as a relational database, it would be possible for a practitioner to, for example, identify all diabetic patients in her practice who are not well controlled (e.g., all patients with HbA1c levels of greater that 6.5%). The physician might then use this knowledge to make a concerted effort to target these patients and get their diabetes under control. If the physician were not sure how to accomplish this, she might seek specific information, making effective use of the self-monitoring process of "knowing when to look it up" (Eva and Regehr 2013). The use of EPRs in this manner highlights the potential of technology to facilitate the professional development of healthcare professionals in a safe and non-threatening environment. The next step for researchers and practitioners is to determine the best ways to use this technology in a way that is both effective and engaging with patients and healthcare providers. This is not an easy task and one that requires careful consideration and planning. However, with the increasing use of EPRs in healthcare, we are likely to see more research in this area in the future.

As a conclusion, we can see that the use of feedback in a constructive manner can lead to positive feelings and improved performance. It is important to note that feedback should be given in a way that is engaging and non-threatening to the patient. The use of technology such as EPRs can be a useful tool in improving performance and patient care. However, it is important to consider the implications of using technology in this way and to ensure that it is used in a way that is effective and engaging for both patients and healthcare providers. It is also important to note that the use of technology in this way is not without challenges and requires careful consideration and planning. It is hoped that this chapter has provided some insights into the use of technology in the healthcare setting and has highlighted the potential for improving performance and patient care through the use of EPRs.
Regehr 2007) to accomplish this specific task. By engaging in this process, the physician will have improved her practice and learned about diabetes management in the process, without ever having had to engage in the summative self-assessment process of deciding she was below average (or below standard) in her ability to manage diabetes. Similarly, informative but relatively nonthreatening data can be derived from peer consultation: Mann et al. (2011) report that physicians involved in peer-oriented professional development groups value the opportunity to learn from the experiences presented in conversation between others without having to directly reconcile the belief that they are practicing well (or the desire to present an impression of practicing well) with information that indicates a gap in one’s practice habits.

As a complicating factor, however, Kluger and van Dijk (2010) have demonstrated, using Higgins’s self-regulation theory (1997), that negative feedback (e.g., failing to perform a task) is generally more efficacious than positive feedback when one is engaged in activity that requires avoiding a negative outcome (prevention), whereas the reverse is true when one is engaged with an activity that requires achieving a positive outcome (promotion). For example, they argue that a surgeon performing a routine procedure will be in prevention focus (i.e., with no real challenge given the surgeon’s skill level, the task is largely one of avoiding mistakes). Success in this orientation is a low arousal condition resulting in little learning or alteration of behavior, whereas failure leads to stress, high arousal, and a greater influence of the (negative) feedback received. In contrast, the surgeon who is trying to perform a very novel and complicated procedure with little chance of success will be rewarded and influenced by feedback much more when a positive outcome occurs than she will be from the frustration that will arise from failing at a task that was not necessarily expected to be successful. In fact, these patterns are epitomized in several aphorisms from the clinical world such as “Big cases, big complications,” which imply that there are times when negative outcomes are not surprising and therefore do not constitute a meaningful source of feedback. Of particular interest for our current purposes, Kluger and van Dijk note that an individual’s orientation in complicated domains like the health professions is anything but static. Across tasks, over the course of time (and, we would add, even within a particular surgery), one’s orientation may sway from performance oriented to prevention oriented and back again.
As such, one cannot assume that any particular feedback delivery strategy will universally improve performance, an argument that is reinforced both by the influence of contextual factors seen in formal feedback intervention studies (Shute 2008) and by recent arguments put forward against algorithmic (and, as a result, often inauthentic) strategies for feedback delivery (Voyer and Pratt 2011).

Consistent with this notion that feedback must be tailored to the specifics of the situation (and, more generally, that context impacts on learning, as outlined by Mylopoulos in chapter 4), a review of the literature on the effectiveness of students’ evaluations of teachers led Marsh and Roche (1997) to conclude that augmenting feedback with consultation (i.e., having someone help the recipient work through the feedback received) is one of the most influential variables with respect to inducing behavior change. These findings could indicate an important distinction between “data” as information we are left to our own devices to interpret and “feedback” as meaning that is co-constructed through social interaction. If we believe Kruger and Dunning’s (1999) argument that being unskilled and unaware go hand in hand, then it follows that counsel from a peer or supervisor may be necessary to help one understand (let alone determine how to use) external data when they are available. Such a counselor can help us place data in context (as guidance on how to improve performance rather than as a statement about our worth as individual practitioners), can minimize the risk of reflexively rejecting data that do not fit our preconceptions by offering alternative interpretations, and can provide a sense of accountability that will make it more difficult to simply ignore pieces of data that might fit uncomfortably within our current worldview.

It is important to note, however, that the credibility of the person giving feedback and the relationship between the person giving feedback and the recipient are likely to be important components of the process. When we interviewed individuals who were selected based on their engagement with formal informed self-assessment activities, we repeatedly heard statements suggesting that perceptions of beneficence and nonmaleficence (i.e., having good intentions towards the recipient) are important determinants of individuals’ receptivity to the feedback provided by others (Eva et al. 2011). That said, the strength of one’s relationship can be a double-edged sword, creating a greater opportunity for the feedback that is delivered to be seen as credible (and delivered in the recipients’ best interest) while also
creating a tension in that a preexisting relationship can make it less likely that feedback will be delivered for fear of damaging the relationship itself (Mann et al. 2011).

Fortunately, guidance exists regarding ways in which situations might be constructed to enable individuals to discover the limits of their knowledge and how performance might be improved without referencing their ability to some norm. Robert Bjork (1999), along with many colleagues (e.g., Schmidt 1991), has put forward the notion that learning takes place predominantly through making mistakes, and that as a result, our task as educators or regulatory authorities is to construct situations that force individuals to encounter the limits of their ability. Termined “desirable difficulties,” this concept is particularly important, as illustrated by Koriat et al. (2004), who have demonstrated that when we are given mistake-inducing tasks that may make us feel less comfortable/confident, we often show greater rates of retention and knowledge transfer than when left to our own devices because we often misjudge the extent to which we are learning (see Boehler et al. 2006; Eva 2009).

Test-enhanced learning offers one example of a strategy for inducing desirable difficulties as a pedagogical intervention, awareness of which is quickly spreading throughout the health professions education community (see Eva 2007; Kromann, Jensen, and Ringsted 2009; Larsen, Butler, and Roediger 2008). “Test-enhanced learning” describes enhanced learning that results from taking a test on to-be-learned material relative to spending the same amount of time restudying that material multiple times. In at least one study, the benefit of taking a test on learning was maintained for six months (Kromann et al. 2010). Greater awareness of this benefit may help inform the efforts of professional development facilitators by leading them to more deliberately deliver formative testing opportunities, thus encouraging greater participation on the part of learners in reviewing and improving their practice through well-constructed self-directed “assessment” activities, such as structured record audits and routine formative examinations.

Greater awareness of the value of desirable difficulties could also extend individuals’ motivation to engage in the habit of self-directed assessment seeking (looking for opportunities to formally test themselves rather than relying on self-perceptions). In fact, while experiencing desirable difficulties has been argued to directly improve the learner’s capacity to recall
information by increasing the strength of the memory for the material, we have recently found that tests may also enhance learning by altering test takers’ tendencies to seek out information of relevance to improving their performance. In a series of studies on self-monitoring (i.e., awareness of one’s likelihood of answering correctly in the moment of being presented with a problem) undergraduate psychology students (Eva and Regehr 2007, 2011) and medical students (Agrawal, Norman, and Eva 2012) were asked to sit for sixty-item computer-based tests of general or medical knowledge, respectively. After completing the test students were four times as likely to search for information relevant to questions that they had to defer answering during the test (due to lack of confidence in their ability to answer the question correctly) relative to questions they chose to answer on the first opportunity (Eva and Regehr 2011). Similarly, they were twice as likely to search for information relevant to test questions if they had to deliberate extensively (>20 seconds) about whether they knew the answer relative to when they could decide quickly (<3 seconds) to answer or defer. Agrawal, Norman, and Eva found that the amount of time spent reviewing answers to individual questions was greater for questions medical students answered incorrectly relative to questions they answered correctly (even though feedback on their personal accuracy was not delivered explicitly). Perhaps more interestingly, however, was that within each of these categories, students were seen to spend more time reviewing questions in which the outcome was discordant with their expectations, spending the most time reviewing questions they answered incorrectly after indicating confidence in their response.

Taken as a whole, the material reviewed in this section generally points to a need to recognize that the recipient of feedback is a much more important component of the feedback process than the feedback intervention literature suggests. The recipient’s perceptions of their own performance, the congruence between those perceptions and the feedback received, and the recipient’s perceptions of the credibility and intent of the source of feedback will all influence the reception of that feedback and, in turn, influence the likelihood of engaging genuine effort toward performance improvement. In fact, we would go so far as to say that at a fundamental level, the recipient is the deciding factor in whether or not change occurs. We are led to this conclusion in part by very recent work indicating that inducing people to think about how others can help them achieve their goals can undermine subjective motivation goals will not lead to change. This report on the role of social influence on self-regulation suggests that social interaction and self-regulation are not isolated, and that we should not evaluate the effects of feedback solely by reference to the changes that occur in our models of self-regulation. Feedback and the role of self-reflection in learning are not mutually exclusive. In fact, learners require a balance of both feedback and understanding of how feedback fits with their prior knowledge to make the most of the feedback. Therefore, we propose that opportunities for feedback to support self-regulation can improve learning by providing an environment that encourages reflective processes.

Conclusion
Self-Assessment and Performance Improvement

...for the material, we achieved by altering test situations to improve their understanding (i.e., awareness of the content) of the presentation and the improvement (i.e., awareness of being presented with questions) (Eva and Regehr 2008, and Eva 2012) of general or medical students.

Students were four questions that they had confidence in their ability to answer, they chose to answer first. Similarly, they were twice as likely to answer questions if they had to answer them. They knew the answer (with confidence) to answer or defer.

The time spent reviewing questions was shorter for answered correctly than for those delivered explicitly. Within each of these categories, reviewing questions in multiple iterations, spending the time immediately after indicating the correct answer generally points to a more important component of the intervention literature on self-regulation, performance, the combination of the feedback received, and the recognition of the source of feedback received, in turn, influence performance improvement. At a fundamental level, the feedback guiding change occurs. We are indicating that inducing individuals’ motivations and thus make it less likely that those goals will be achieved. In a series of studies Fitzsimmons and Finkel (2011) report on the impact of “outsourcing” self-regulation. While they note that social support can benefit individuals in their pursuit of goals (Rusbult, Finkel, and Kumashiro 2009), they found that participants who were encouraged to think about how their partners could help them achieve health or academic goals were more likely to procrastinate and less likely to demonstrate behavior that would help them achieve those goals over the course of the following week relative to control groups of participants.

The change we hope to induce at the community level, therefore, is not one of eliminating the encouragement of self-directed learning. Rather the change we hope to induce is one of greater awareness that best practice in this area is not defined by the accuracy of one’s self-assessment, that accurate self-assessment is not the mechanism whereby performance is improved, and that we should not judge the effectiveness of educational interventions solely by reliance on the perceptions of the learners. Instead, we should alter our models of self-directed learning to focus on encouraging self-directed assessment seeking. Doing so will involve shifting the conversation around the role of self-assessment toward the idea that the most important teaching we can do with respect to nurturing “good self-assessors” is to help learners recognize with humility that we are not uniquely privileged in understanding the strengths and limits of our own behaviors. Models of continuing professional development should, therefore, focus on creating opportunities for individuals to discover ways in which their performance can improve without prioritizing accuracy of self-assessment as the mechanism through which such improvements take place.

Conclusions

Irony is a disciplinarian, feared only by those who do not know it, but cherished by those who do.

KIERKEGAARD, The Concept of Irony

Continuing our appreciation of irony, throughout this chapter we have argued that we cannot trust self-assessment but should encourage it; that we need to influence self-assessment but should not change it; and that we can best improve the accuracy of self-assessment by ignoring it. We
appreciate that these are difficult concepts, but the practical implication is that self-assessment does have a major role to play in performance improvement so long as it is not treated as the mechanism through which performance improves. When self-assessment is made the focus of individuals’ determinations of what activities should be engaged to yield improved performance, it acts as a powerful lens through which all data and feedback are interpreted. All the research we can find suggests that the result of these processes is an interference with practice improvement, not its enhancement. Indeed, our judgments about ourselves drive as much as they reflect performance, but in ways that do not always appear rational. Helping poorer performers better understand that they lie in the lower quartile of their cohort is likely to be counterproductive. Such an understanding does not encourage them to improve, but rather discourages them from even trying. Thus the path to improving practice likely does not lie in helping people improve their self-assessment ability. Rather, it is more likely to lie in focusing instead on creating desirable difficulties that enable individuals to struggle effectively with their performance, prompting them to make mistakes from which they can learn. Ironically, such a process will also lead to the appearance of improved self-assessment by raising performance to meet our optimism rather than improving our actual self-assessment. Better that, however, than diminishing our optimism to the point that performance will not be raised. We are blinded by the insights provided through our self-assessment, but that is something to be taken advantage of, rather than something to be overcome.

Judgment has conceptualized the role of the heads of physical performance in medicine. The surgeon can best detect a patient’s core competencies by analyzing the seal of medical practitioners in psychology in health. Realizations of this are put out on a daily basis in a vacuum. Therefore, the role of individuals becomes decision-making.