



The Debate on Educational and Psychological Testing in the United States:
An Essay Review and Outsider Perspective with Some Philosophical Musings

Jesper Garsdal
&
Christian Ydesen
University of Aarhus

Madaus, George; Russell, Michael; and Higgins, Jennifer. (2009). *The Paradoxes of High Stakes Testing*. Charlotte, NC: Information Age Publishing.

Pp. 264

ISBN 978-1-60752-027-6

Phelps, Richard P. (Ed.) (2009) *Correcting Fallacies about Educational and Psychological Testing*. Washington, DC: American Psychological Association.

Pp. 287

ISBN 978-1433803925

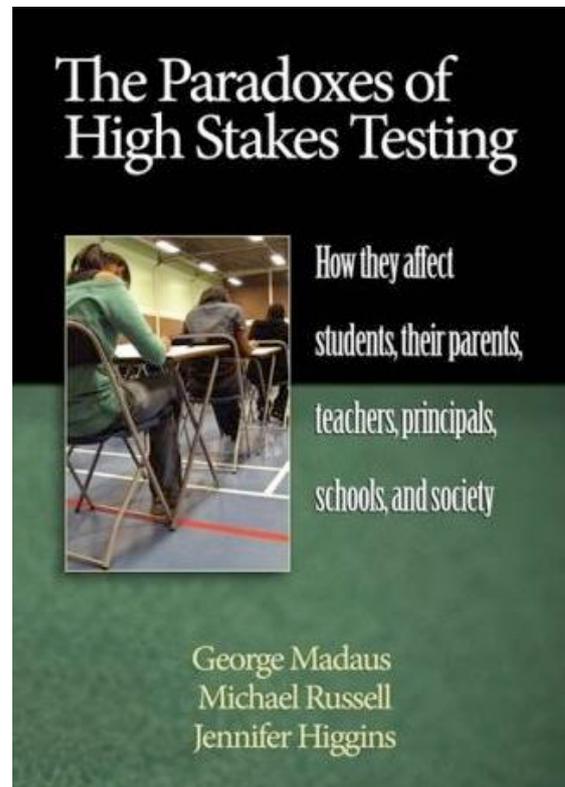
Citation: Garsdal, Jesper & Ydesen, Christian. (2009, June 15). The debate on educational and psychological testing in the United States: An essay review and outsider perspective with some philosophical musings. *Education Review*, 12(8). Retrieved [date] from <http://edrev.asu.edu/essays/v12n8index.html>

The year 2009 has witnessed the publication of two canonical but very different works on educational and psychological testing: Madaus, Russell and Higgins's *The Paradoxes of High Stakes Testing* and Phelps's *Correcting Fallacies about Educational and Psychological Testing*. Both books are written by experienced and esteemed researchers of educational and psychological testing, but even so the books clearly testify to the ongoing—and sometime acrimonious—debate on testing in the United States. This is not to say that the books are substandard but merely that their approaches and conclusions do not fall short of being antagonistic. In fact, from a philosophical analytical perspective, the books reveal two different methodological paradigms for addressing the phenomenon of educational and psychological testing: an empirical hypothetical-deductive approach based on objective measurement (the Phelps anthology) and a sociological approach incorporating the human, cultural, historical, and

technical dimensions of testing (Madaus, Russell, and Higgins). This difference of paradigm is indeed telling since testing comprises numerous dimensions relevant to different research disciplines with their diverging methodologies.

However, the two works share an ambition to analyze different arguments about testing, but it is no exaggeration to say that they draw irreconcilable pictures of the contemporary testing climate in the US.

The Phelps anthology paints a picture of a division between scientists and advocates in the field of testing (Phelps & Gottfredson, 2009, pp. 250ff.). The book is written with a sense that the testing community has been severely wronged by a powerful group of sophists ignoring the scientific facts of testing. The authors launch a systematic broadside against a vast array of test critics (Phelps, pp. 96ff.) (Phelps & Gottfredson, 2009, pp. 247ff.). In the words of Linda Gottfredson: "Mere ignorance of the facts cannot explain why accepted opinion tends to be opposite the experts' judgments. Such opinion reflects systematic misinformation, not lack of information. The puzzle, then, is to understand how the empirical truths about testing are made to seem false, and false criticisms made to seem true. In the millennia-old field of rhetoric (verbal persuasion), this question falls under the broad rubric of sophistry." (Gottfredson, 2009, p. 18). The ambition of the Phelps anthology is simply to rehabilitate and promote educational and psychological testing.



George Madaus

Madaus, Russell, and Higgins, on the other hand, point to a number of problematic aspects in current U.S. testing practice often promoted by eager politicians advocating tests. They seek to draw a holistic picture of testing as a societal phenomenon and to raise awareness of the testing arguments and practices prevalent in the U.S. While they do not argue for the abolition of tests, they advocate a critical and democratic employment of tests (Madaus, Russell, & Higgins, 2009, p. 3 & pp. 197ff.).

Writing a comparative essay review of these irreconcilable works could very well benefit from an outsider perspective—not the least because of the politicized nature of the debate on testing which poses a serious problem for science and scholarship in the field. As two Danish philosophers, we will try to raise some thought provoking issues aimed at putting the

antagonism into perspective. One way of approaching the antagonism between the two books is to establish an expanded and constructive plateau of observation consisting of mutual categories and transversal concepts inscribed as a third party of dialogue. This plateau might be sustained by four pillars; empirical data, scientific logic, philosophy, and interculturality, each referenced to some central arguments and dimensions covered in the two works.

Empirical Data—The Question of Intelligence

In regard to empirical data about intelligence, the Phelps anthology states that intelligence tests measure a general intelligence in the individual (Gottfredson, 2009, p. 30). Madaus, Russell, and Higgins also address the question of a general intelligence, but they are more focused on the different historical notions of intelligence and the fact that intelligence is an inferred phenomenon (Madaus, Russell, & Higgins, 2009, p. 64 & pp. 122ff.), a fact which by the way is supported in the Phelps anthology (Gottfredson, 2009, p. 19).

However, while the Phelps anthology clearly recognises the validity and reliability of intelligence tests, Madaus, Russell, and Higgins are more sceptical. They describe a wide array of error types in testing, and in their chapter on the history of testing they state that: “The belief that ‘intelligence’ test scores reflected innate ability ignored the fact that Binet had developed his tests to measure developed ability, rather than innate ability. Binet developed his test to identify students in need of specialized instruction. As the Stanford psychologist Richard Snow put it, ‘to interpret [Binet’s test] as measures of “general intelligence” was a flagrant over generalization.’ This overgeneralization, however, has persisted since its inception a century ago.” (Madaus, Russell, & Higgins, 2009, p. 123). Thus, the question of general intelligence clearly distinguishes the two books.



Michael Russell

The big issue about the question of intelligence seems to be whether the empirical datasets prove the existence of a general intelligence or if general intelligence is merely a historical, geographical, and sociological construct subject to local constructions of meaning. Or as Madaus, Russell, and Higgins write: “(...) all interpretations and descriptions in science are normative.” (Madaus, Russell, & Higgins, 2009, p. 125).

It certainly is not the ambition of this essay review to resolve this matter but merely to expand the field of debate pertaining to the two books.

What seems to be missing in the work by Madaus, Russell, and Higgins is a clear and upfront discussion of the vast empirical datasets and previous studies on intelligence testing relating them to notions of evidence, significance, practice, and theory of science. On the other hand, what seems to be missing in the Phelps anthology is some in-depth critical reflections on the relevant categories employed when dealing with intelligence testing—races, groups, sub-domains, heritability. Lack of such reflections might lead to a barren reproduction of problematic categories. In other words, employing the same sort of instrument generates the

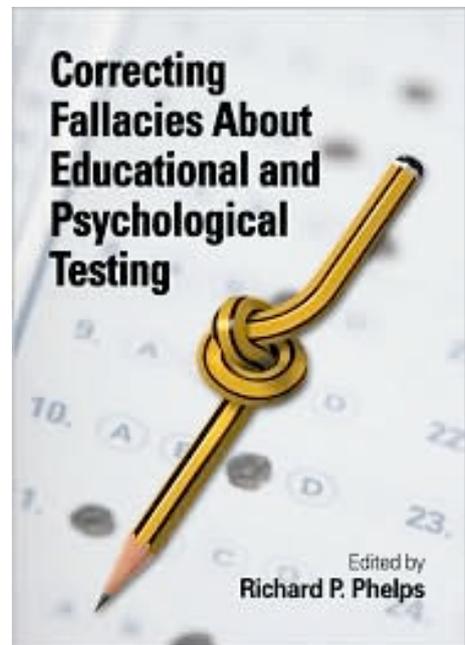
same kinds of answer. But that does not prove the validity of the instruments because they are based on the same assumptions and hypotheses regarding correlations, causal effects, society, and human nature. In fact, arguing for an empiricist-metaphysical notion of intelligence as “real” calls for a philosophical discussion touching on other notions of intelligence in order to avoid narrow and closed self-reference and self-justification. Otherwise it would seem that we slide back into the unresolved problems of logical positivism, which rests on a metaphysical claim that there is no metaphysics, only empirical data and logic. Thus, a claim about the existence of general intelligence is therefore in itself problematic. It is not sufficient to just refer to statistical correlations. The issue must be opened up to philosophical discussion.

However, what seems to be more or less missing from both works is the notion that concepts are neither only empirically nor culturally given, they also have philosophical backgrounds, connotations, and implications, as for example the notion of intelligence. Madaus, Russell, and Higgins do however discuss the importance of “names” in regard to testing (Madaus, Russell, & Higgins, 2009, p. 64f.); but we should also emphasize the importance of borrowing connotations from concepts developed in the history of ideas. We will expand this point in the discussion of interculturality; but for the time being, suffice it to say that it is not adequate to just unleash a scientific definition of general intelligence without having addressed the conceptual level.

Scientific Logic—Description or Understanding

This brings us to the second relevant pillar for discussion, namely, scientific logic, since both works recite the mantras of science. The Phelps anthology in particular makes a very strong case about logic since it tries to use logic as a way of falsifying the arguments put forth by various test critics (Phelps, p. 4) (Gottfredson, 2009) (Phelps & Gottfredson, 2009).

But while the Phelps anthology seeks to blow contrary and critical positions out of the water, Madaus, Russell, and Higgins simply ignore the positions of the Phelps anthology, or they are at least not unfolded in any manifest way. Such is often the strategy of test critics, a strategy often condemned by Phelps and his contributors (Gottfredson, 2009, p. 58) (Phelps, p. 91) (Camara, 2009, p. 174) (O'Boyle Jr. & McDaniel, 2009). Madaus, Russell, and Higgins attempt to transcend the strict scientific logic of measurement and quantification through historical and sociological analysis combined with an analysis of the underlying assumptions of the policy of test promotion in American education (Madaus, Russell, & Higgins, 2009, p. 13ff.). They reach a preliminary conclusion that “For proponents of high-stakes testing, there is both a willed ignorance and an imperious immediacy of interest that leads proponents to turn a blind eye to the fallibility of test scores, and the paradoxical, chronic negative consequences of using test scores to make high-stakes decisions.”



They continue: "... relying on the same single quantitative measure for all students and schools as the key reform tool is analogous to using only a sledge hammer to renovate a house." (Madaus, Russell, & Higgins, 2009, pp. 22-23).

This brief view on the scientific logic of the two works calls the old scientific distinction between description and understanding to mind. Thus, it is reasonable to observe that the books testify to a clash between a binary logic of empiricism and a more open logic of contextualisation, embeddedness, and interpretation.

Seen from our philosophical plateau of observation, logic can be a useful tool in a scientific process but it is not exhaustive because it is related and interdependent with problems in scientific theory. There is not only different cultural "logics," but equally important there are also conceptual differences in different philosophical logics about what "inferences" and "concepts" "really are. A claim about sticking to science does not solve this fundamental problem. Blowing opponents out of the water using logic as a weapon might be an understandable endeavour, keeping the methodological paradigm of the Phelps anthology in mind, but from a philosophical point of view it amounts to no more than raising a voice from an imagined no-man's land seeking to uniform and patent the discussion without any room for dialogue. Instead a degree of humility towards the production of knowledge might be more becoming because it recognises that one's own logic does not negate the meaning inscribed in a contrary position. Different language games make use of different kinds of logic.

Philosophy—Quality, Quantity, and the Individual

Inferring a brief characterization of testing from the two books in combination would imply that testing is tantamount to the transformation of quality into quantity with the purpose of screening, selecting, and placement of individuals (Madaus, Russell, & Higgins, 2009, p. 116f.) (Gottfredson, 2009, p. 12). Thus, the notion of the individual involved in testing is extremely important to discuss.

Testing necessarily conceptualizes and anticipates the existence of some kind of invariant and/or path-dependent unique characteristic—intelligence or a structure of ability—in the individual—i.e. an essence—which can be identified accurately in relation to a pre-constructed notion of a knowledge domain (Gottfredson, 2009, p. 16).¹ In other words, testing converts

¹ The British philosopher John Stuart Mill (1806-1873) made a very good point which certainly is thought provoking in relation to both the notion of an essence in the individual and the notion of knowledge domains: "The tendency has always been strong to believe that whatever received a name must be an entity or thing, having an independent existence of its own; and if no real entity answering to the name could be found, men did not for that reason suppose that none existed, but imagined that it was something peculiarly abstruse and mysterious, too high to be an object of sense." (Mill, 1869, p. 5).

abstract epistemological concepts into ontological entities. Testing is directed towards the *individual* test taker, and it is directed towards the future, and in order to be justified must claim a certain level of invariance in the individual test taker—otherwise testing would be futile.

This poses a metaphysical question about the nature of this essence and more broadly what constitutes the individual. The question has sparked a long trail of controversy in the field of testing about the role of nature and nurture. This is a dividing line which is also applicable to the two books discussed here. Where the Phelps anthology emphasizes the significance of nature (heritability) in testing (Gottfredson, 2009, p. 36ff.) Madaus, Russell, and Higgins must be characterized as leaning more towards nurture as the main explanatory factor (Madaus, Russell, & Higgins, 2009, pp. 61ff.). But what about the possibilities of testing to “capture” the individual—in philosophical terms the transformation of quality into quantity in the process of testing?

Sadly, this question is only addressed indirectly in the Phelps anthology. Although Camara expresses the reservation that “Standardized tests and any metric used for selection or accountability cannot capture the whole breadth of factors that will ultimately determine future success” (Camara, 2009, p. 154), the general picture of the Phelps anthology is that although testing has limitations the empirical evidence of correlations proves the reliability, validity, and fairness of testing (Gottfredson, 2009, p. 14) (O’Boyle Jr. & McDaniel, 2009, p. 182ff.) In other words, quality can be adequately quantified by a meticulously designed test. But is it really possible to reduce the individual to a cocktail of numbers? Can a focus on pre-constructed samples from an equally pre-constructed domain give reassurance that nothing has evaded its view? Madaus, Russell, and Higgins have a very well-presented chapter explaining this transformation of quality into quantity through a process of sampling from a domain in the process of constructing a test (Madaus, Russell, & Higgins, 2009, pp. 37ff.). They argue convincingly that testing by its very nature is often at risk of systematising and denying the test takers their individuality because they must live up to the pre-constructed logic of a one-size-fits-all test (Madaus, Russell, & Higgins, 2009, p. 23). This is so even though tests are standardised and administered under the same conditions for all test takers. They argue that test takers inevitably are different and therefore a standard test cannot incorporate their differences. Instead, the test is in imminent risk of stifling and suspending idiosyncrasies. It is important to note that the test taker’s score is dependent on both his/her performance and performance of the norm-group to which he/she is compared.

Thus, the big issue with testing is that it seeks to quantify, retain or capture something qualitative in order to produce comparable and standardised results. If we try to follow that logic of quantifying quality it becomes apparent that a quantum can be so big that it transforms qualitatively. This implies that another yardstick is needed in order to measure the new quality. The consequence of this observation is that quantity or measurement becomes relative because a yardstick is needed to measure the other yardstick and so on; this generates an infinite regress (Himmelstrup, 1964, p. 208). The reason is that at least part of quality always seems to escape quantification attempts. The great German philosopher Georg Wilhelm Friedrich Hegel (1770-1831) writes about this problem of quantification: „*Ihr Zweck oder Begriff ist die Größe. Dies*

ist gerade das unwesentliche, begrifflose Verhältnis. Die Bewegung des Wissens geht darum auf der Oberfläche vor, berührt nicht die Sache selbst, nicht das Wesen oder den Begriff, und ist deswegen kein Begreifen.“ (Hegel, 1988, p. 33). The quotation from Hegel makes the somewhat provocative statement that quantity is an insignificant difference. The reason is that quality with logical necessity is prior to quantity—it is impossible to have a quantity that is not a quantity of something. The point is notably supported by the Danish statistician of educational testing Georg Rasch (1901-1980) who used to say to his students that all research is at its bottom qualitative because quantitative research is always a quantification of something qualitative (Allerup, 2005).

Philosophically it can be said that it is impossible to describe an individual qua individual—the test can only describe the test taker in relation to other test takers or a predetermined standard. This amounts to a paradox about testing not fully developed by Madaus, Russel, and Higgins: On the one hand testing denies the test taker his/her uniqueness but on the other hand testing is a science of the individual. Thus, testing is bound to disclose the uniqueness of the test taker which is the very thing that it fails to measure. Instead testing can only relate to itself and therefore the results of a test will always be chained to a perspective from its own totality circularly confirming its own righteousness.

Democracy, Human Dignity and Interculturality

What is particularly interesting about testing is that it forms an entry for disclosing the social ontology of society. For instance, educational testing from its very dawn often subscribed to meritocratic and emancipatory ideals: the notion of the poor, intelligent child (Oakland, 2009, pp. ix ff.) (Madaus, Russell, & Higgins, 2009, p. 123). Both works include sporadic discussions about this link between testing and social ontology, particularly in regard to notions of democracy. In this final pillar of relevance we will attempt to take the perspective of the other—an intercultural perspective—in relation to democracy and the broader question of human dignity in testing.

Madaus, Russel, and Higgins dedicate a chapter to a discussion of human and cultural factors in regard to testing. They point out that: “High-stakes testing incorporates two culturally held values. The first is that achievement is an individual accomplishment. The second value is that individuals must display their accomplishment publicly.” (Madaus, Russell, & Higgins, 2009, p. 62). And they continue: “Tests tend to promote the values of objectivity, the importance of factual knowledge and ‘right’ answers, and rapid visible performance. In so doing, tests devalue subjectivity, feelings, reflection, introspection, and discernment.” (Madaus, Russell, & Higgins, 2009, p. 63). Contrary to this position, the Phelps anthology states that there is no cultural bias in testing (O’Boyle Jr. & McDaniel, 2009) (Camara, 2009, pp. 158 ff.) (Gottfredson, 2009), that the validity of the bell-curve is beyond reproach, and finally that the bell curve of phenotypic intelligence for Asians and Jews is slightly higher than for whites which again is slightly higher than the one for blacks (Gottfredson, 2009)(see also appendix B). There simply is no attention paid to the existence of immanent values of a test in the Phelps anthology as they are deemed irrelevant. There is, however, a notion that testing can serve “... some social interests and goals

over others” (Gottfredson, 2009, p. 52) but that does not mean that something is wrong with the test, only that the context in which it is employed is wrong. This is a big difference in relation to Madaus, Russell, and Higgins who hold that something also might be wrong with the test itself.

This antagonism becomes particularly evident in relation to the notion of democracy. Madaus, Russel, and Higgins call for the implementation of an independent monitoring body of tests in the name of transparency and democracy (Madaus, Russell, & Higgins, 2009, pp. 197 ff.). This call clearly reflects an underlying assumption with Madaus, Russell, and Higgins that testing is a mere tool for society—a necessary but far from perfect tool containing numerous measurement and human errors.

The Phelps anthology evidences a quite different approach to testing, stating that even though testing is not perfect it still reveals reality and society should act accordingly (Phelps, 2009a, p. 3) (Gottfredson, 2009, p. 52). In fact, it is stated that because individuals differ substantially in merit (as demonstrated by tests) democratic societies cannot satisfy both equal opportunities and equal outcomes (Gottfredson, 2009, p. 13). In other words, inequalities in phenotypic intelligence levels generate problems for democracy. But is this really a problem? Limiting the discussion to a question of phenotypic traits and two apparently incommensurable ideals of democracy seems amputated. Instead it would have been much more interesting to focus on another very crucial democratic ideal which is sadly absent from the discussions in the Phelps anthology, namely the notion that all persons are created equal. Following the arguments of the anthology, it is relevant to ask if the findings of testing about the phenotypic inequalities among people are compatible with democracy at all.

Describing society as a cocktail of phenotypes and genotypes seems to be a farfetched and reductionist notion which is unable to take the notion of human dignity into account. Since Pico della Mirandola (1463-1494), dignity has played a central part in humanistic approaches to “Western” education. Naturally, this notion of human dignity implies certain ideas about what it means to be human and it can of course be challenged and subject to debate from philosophical and intercultural perspectives. But what is not acceptable is to introduce a philosophically primitive notion of the individual as some sort of cocktail between genotype and phenotype, without relating to other complex notions of individuality, or to avoid this debate by making claims about sticking to science. Such a position at best demonstrates a certain conceptual naiveté regarding problems of philosophical anthropology.

Finally, as mentioned earlier, we will also emphasize the importance of an awareness of connotations of philosophical concepts developed in the cross-cultural history of ideas regarding the core notions in testing, as for example the idea of intelligence. A good example of such intercultural exchange and transformation is the way notions of active and passive intelligence altered in medieval times in the interaction between Greek, Jewish, Islamic, and Christian philosophy. Several ways of *conceiving* these two aspects of intelligence were introduced and related to different religious paradigms and interpretation of these paradigms, but at the same time they were also part of a lively cross-cultural discussion about the idea of intelligence which took place across cultural and religious borders. Conceptual reflections on

such a distinction and the implications of this intercultural conceptual reflection on the notion of intelligence are not unimportant, just because they are old. The focus on empirical evidence for a general intelligence can to a large extent be seen as born out of a nominalistic position (also developed in medieval times) regarding the way we should conceive term and concepts, but this position is in no way the only stance one can take towards these issues. For example the distinction between active and passive intelligence has been central for the development of Peircian logic and semiotics, a fact Charles Peirce (1839-1914) himself acknowledged by referring to Duns Scotus (c. 1266-1308) (and indirectly to Thomas of Erfurt). Such considerations might not make testing methodology easier, but it might raise an (intercultural) awareness of the implicit metaphysical assumptions and the tendency to reduce and monopolizing a complex notion such as intelligence into a very narrow framework. Furthermore, it might also help to bring the discussion about testing up at a more satisfying conceptual level, where it is recognized that scientific definitions are not the same as concepts.

Concluding Remarks

As we have demonstrated, the two works testify to two very different paradigms regarding their methodological approach to the phenomenon of testing. Both approaches are valid but as we have tried to demonstrate, a number of fundamental questions regarding theory of science, philosophy, and conceptual awareness remain unaddressed. This is not to call for the inclusion of a doctoral dissertation in philosophy in both works but just to say that both books might benefit and strengthen their arguments by including some philosophical reflections on evidence, significance, practice, theory of science, and not least, concepts.

Moreover, both books seem to aspire to make a difference on the political level, each trying to sway public opinion and perhaps the opinions of policy makers about testing. But does that mean that the books do not steer completely clear of a political agenda in their communication of research? The Phelps anthology contains sentences like “With the election of George W. Bush, GOP [Republican] policy advisors faced a historic opportunity, with enormous implications, to benefit U.S. education.” (Phelps, p. 115). Madaus, Russell, and Higgins do not make that kind of statement. They are generally critical towards politicians promoting testing in the field of education. They are explicitly critical of the No child Left Behind (NCLB) law; but at the same time they emphasize the collaboration of both Republicans and Democrats in the passing of the act (Madaus, Russell, & Higgins, 2009, p. 8). In fact, they dedicate a paragraph to the question of why “... test-based accountability proposals resonate across the political and ideological spectrum, left, right, and center.” (Madaus, Russell, & Higgins, 2009, pp. 28 ff.).

However, the bottom line is that both books are exceptionally well written, and they certainly testify to a very high level of communicative skill. The worst shortcoming of both books from an outsider’s perspective is their inability to engage properly each other’s position on multiple levels. This heightens the risk of just reproducing their own logic and discourses as well as ultimately their very own research through exchanges within a like-minded community. Madaus, Russell, and Higgins do not address the main points of the Phelps anthology about the huge empirical evidence allegedly supporting the practice of testing. The Phelps anthology, on

the other hand, is permeated with an aura of indignation against test critics leading to a very binary portrait of the testing climate in the U.S., maintaining that test critics are generally wrong and the testing community is generally right (Phelps & Gottfredson, 2009, p. 247). Two examples are the briefly mentioned dismissal of the research of Stephen J. Gould and the rehabilitation of Cyril Burt (Gottfredson, 2009, p. 55), although the anthology does also contain a couple of more varied accounts (Camara, 2009, pp. 153 ff.).

As outsiders to the field of testing in the United States wanting to learn more, we would highly recommend reading both books as a supplement to each other. It might be a confusing experience, but the confusion will definitely be at a higher level.

Nevertheless, in closing we will mention a Sufi story to illustrate, what the field of testing might end up with if some of the above mentioned comments are not taken into consideration. A friend of the well known Sufi Nasruddin one day found Nasruddin crawling around outside his house. Nasruddin explained that he had lost a coin, and the friend crawled around for a long time with him, where they both tried to find the coin. They did not find it and finally the friend asked, if Nasruddin could remember where he had lost the coin more precisely. "I lost it in the house" was Nasruddin's answer. The friend jumped up and cried "but why are we looking for it here then?"; Nasruddin answered "it is easier to find something, where there is light."

One might wonder if the search for equal opportunities, the individual, and general intelligence might not be better off if the search was done elsewhere than in the light of the political-economical paradigm and control ideology of testing?

References

- Allerup, P. (2005). Hvad er evidens i uddannelser og i professioner? *Agora*, 7, 11-14.
- Camara, W. J. (2009). College Admission Testing. In *Correcting Fallacies About Educational and Psychological Testing* (pp. 147-180). Washington DC: American Psychological Association.
- Gottfredson, L. S. (2009). Logical Fallacies Used to Dismiss the Evidence on Intelligence Testing. In R. P. Phelps (Ed.), *Correcting Fallacies about Educational and Psychological Testing* (pp. 11-65). Washington DC: American Psychological Association.
- Hegel, G. (1988). *Phänomenologie des Geistes*. Hamburg.
- Himmelstrup, J. (1964). *Terminologisk ordbog til Søren Kierkegaards Samlede værker*. Copenhagen: Gyldendal.
- Madaus, G., Russell, M., & Higgins, J. (2009). *The Paradoxes of High Stakes Testing*. Charlotte: Information Age Publishing.

Mill, J. S. (1869). *Analysis of the Phenomena of the Human Mind* (2 ed.). London: Longmans, Green and Dyer.

Oakland, T. (2009). Foreword. In *Correcting Fallacies About Educational and Psychological Testing* (pp. ix-xii). Washington DC: American Psychological Association.

O'Boyle Jr., E. H., & McDaniel, M. A. (2009). Criticisms of Employment Testing: A Commentary. In *Correcting Fallacies About Educational and Psychological Testing* (pp. 181-197). Washington DC: American Psychological Association.

Phelps, R. P., & Gottfredson, L. S. (2009). Summary and Discussion. In R. P. Phelps (Ed.), *Correcting Fallacies about Educational and Psychological Testing* (pp. 247-255). Washington: American Psychological Association.

About the Reviewers

Jesper Garsdal, Ph.D. in Philosophy, is Assistant Professor at Aarhus College of Education, Via University College, Denmark and coordinator of the MINT-network (Meetings on Intercultural Thought). His research interests include intercultural philosophy and history of ideas, philosophy of education, philosophical anthropology, religious studies and community psychology.



Christian Ydesen is a doctoral student in history of education at the Danish school of Education, University of Aarhus, Denmark. His research interests include intercultural philosophy and education, theory of science, philosophy of education, and the history of educational testing.





**Copyright is retained by the first or sole author,
who grants right of first publication to the
Education Review.**

Editors

**Gene V Glass
Arizona State University**

**Gustavo Fischman
Arizona State University**

**Melissa Cast-Brede
University of Nebraska, Omaha**

<http://edrev.asu.edu>