

Non-cognitive factors in health sciences education: from the clinic floor to the cutting room floor

Geoff Norman

Published online: 19 November 2009
© Springer Science+Business Media B.V. 2009

This editorial represents a considerable departure from tradition. In the Reflections section of this issue, Lane has written a comprehensive review of literature related to “Professional Competencies”. The editorial is a response to the review, and is written with the knowledge and approval of Dr. Lane. It was prompted by my editorial review of her submission, and is intended to provide an opportunity for a fuller exploration of the whole domain of non-cognitive skills. In past issues, we have had some critical reviews of some of the areas touched on by Lane (Lewis et al. 2005; Eva 2005; Norman 2008), but the very wide range of topics she considers provides, in its turn, an opportunity for a similarly comprehensive critique. As the editorial is a direct response to the Reflections article, readers are encouraged to read the Reflections article first, then return to the editorial.

Lane’s review is wide ranging, bringing in subjects as disparate as emotional intelligence, personality, learning style, and technical skills. The review is vast; in the preprint version, the reference list extended to 20 pages. One indication of the breadth of the review is revealed by examining the various labels she uses to describe these competencies:

non-technical, ethical, interpersonal, self-management, leadership, business, thinking, bedside manner, non-cognitive variables, humanistic.

And that’s just the abstract

Intuition, art, practical knowledge, know-how, artistry, insight, judgment, connoisseurship, tacit knowledge, implicit knowledge, interpersonal interactions, communication ability.

And that’s page 1.

While the number of labels appears overwhelming, I sense that this diversity is a consequence of the nature of the topic as much as anything else. The field is broad and confusing, and the review inevitably delves into many domains.

Let’s start with the much-abused term, “non-cognitive”. Literally this means “unthinking.” No wonder we have a problem; mounting a program to teach or assess

G. Norman (✉)
MDCL 3519, McMaster University, 1200 Main St. W., Hamilton, ON L8N3Z5, Canada
e-mail: norman@mcmaster.ca

unthinking skills cannot be easy. But that's not what we mean, of course. We don't assess a person's ability to not think. Instead, we call it "non-cognitive" to distinguish it from "cognitive".

But what do we mean by cognitive? I expect, we are referring to the formal, analytical, logical knowledge of the basic and clinical science courses; the physiological mechanisms, the anatomical bits, and the lists of signs and symptoms of the clinical disciplines. In short, knowledge, usually verbal, that can be put on a page (or on a computer screen).

Non-cognitive is presumably what's left over. But as the examples above illustrate, there's a lot left over. Many of the activities under this rubric, such as effective communications or skilful use of instrumentation, require much cognitive activity. After all, even sex, when performed well, has a large cognitive component. It would be really helpful to find a better word, or even none at all, as the lumping of all these various bits under the non-cognitive umbrella imposes an illusion of unity that does not hold up to scrutiny. It's not just an issue of language precision; by grouping all these various knowledges, skills and attitudes under a single umbrella, we virtually guarantee a confusion in research methods and results. As a simple example, suppose we set out, as many admissions committees do, to develop or locate a good measure of "non-cognitive skills". Can we really presume that any one measure will be able to, at one stroke, capture all the personal and developmental aspects that would lead one person to be better at communicating, accessing tacit knowledge, holding a scalpel, and making ethical decisions? Of course not. But that is the Herculean challenge underlying any attempt to assess or select "non-cognitive skills". It is not surprising, therefore, that Lane has wandered far and wide through a number of domains to identify areas of research that may help to identify potentially valuable ways to enhance these skills. To some degree, I will be forced into the same far-reaching approach, simply to respond. But let me attempt some demarcation.

Tacit knowledge

Lane's first dichotomy is between the "art" and "science" of medicine. She equates the "art" of medicine with experiential tacit or implicit knowledge. But while Patel et al. (1999) implied that such cannot be articulated and hence investigated, much has happened in the past decade to reveal many aspects of this knowledge. Whether it is associated with "dual processing" (Evans 2008), "non-analytical" reasoning (Norman et al. 2007) or "exemplars" (Brooks 1978) we now understand, even down to the physiological level, that the brain can process information in at least two fundamentally different ways; one explicit, logical and conceptual, and the other implicit, automatic, beyond introspection, and contextual. But the important part is that both processes can be and are used to achieve a variety of goals, in both "cognitive" and "non-cognitive" domains. Pattern recognition processing, although based on access to "tacit knowledge" is inherent in the highly cognitive act of arriving at a diagnosis, (and many other categorization and problem-solving activities) hence cannot be equated with communication, interpersonal relations etc.

Practical knowledge and learning style

In the review, Lane makes an easy transition from tacit, experiential knowledge to the practical workplace setting. Such an association is quite reasonable; we are, after all, referring to knowledge associated with specific experiences. It would seem to be natural to

then move to an examination of individual differences in practical abilities, inclinations, styles and so forth. And the Kolb et al. (1999) learning style inventory (LSI) does indeed explicitly identify one subgroup as practical. Similarly, champions of multiple intelligences, such as Gardner and Sternberg also delight in placing practical intelligence in opposition to theoretical or abstract intelligence. Taken at face value, these research programs appear ideally suited to provide real value added in the quest to educate better professionals.

Unfortunately, these are two areas that simply should not be taken at face value. Both have attracted, vocal, even vitriolic, opposition. Both appear characterized by a loose interpretation of evidence.

First, learning style. There have been a number of serious challenges to the Kolb approach. Reynolds (1997) reviews studies showing that it is unreliable and invalid (p. 117), and the central premise, that matching learning style to teaching method will enhance learning, lacks any support. As Joughin (1992) puts it, this ignores “both the potential value of creatively mismatching teacher and learner *and the equivocal outcomes of research on matching itself (italics mine)*”. Similar conclusions have been drawn by Coffield et al. (2004), whose comprehensive review concludes with:

“Moreover the status of research in this field is not helped by the overblown claims of some of the developers and their enthusiastic devotees.”

Results in medical education are consistent with these critiques. David Cook has conducted several studies of learning style. One study (Cook et al. 2007) showed no interaction between active or reflective learning style and instruction in predicting performance outcomes. Cook et al. (2009) used a learning style questionnaire similar to Kolb’s, and showed no interaction between learning style and type of instruction, or preference for type. In short, we have yet to see any consistent evidence that learning style is related to preference for instruction or that matching learning style to instruction (the original promise) leads to gains in amount of learning (Norman 2008).

Multiple intelligences practical intelligence, and emotional intelligence

The focus on maximizing practical, instead of theoretical, learning also leads Lane into various aspects of multiple intelligences. This is an old tradition in psychology, championed by two individuals—Howard Gardner and Robert Sternberg. Gardner is more of a theoretician and populist, but Sternberg is a highly respected and credentialed psychologist, chair of psychology at Yale and past president of the American Psychological Association. As Lane describes, Sternberg’s thesis is a “triarchic” theory of intelligence, with three major divisions. It is often not a simple task to discern from Sternberg’s writing precisely what the three divisions, and their various subdivisions, are. For example (Sternberg 1984):

“The first relates intelligence to the internal world of the individual, specifying the mental mechanisms that lead to more or less intelligent behavior.... The second subtheory specifies those points along the continuum of one’s experience with tasks or situations that most critically involve the use of intelligence... The third subtheory relates intelligence to the external world of the individual, specifying three classes of acts—environmental adaptation, selection and shaping—that characterizes intelligent behavior in the everyday world.” (p. 5)

However, one easier interpretation is that the three forms are analytical, practical, and creative (Sternberg 2000). Whatever the theory, Sternberg's central claim is a separation between academic and practical intelligence, and that

“[practical intelligence] is at least as good a predictor of future success as the academic form of intelligence that is commonly used in so-called tests of general intelligence [g]. Arguably, practical intelligence is a better predictor of success.” (Sternberg et al. 2000a, pp. xi–xii.)

Thus central to the Sternberg claim is that we should expand our view of intelligence to consider measures of practical intelligence, which has much greater relevance to real world tasks. Sternberg again:

“... we argue that practical intelligence is a construct that is distinct from general intelligence, and that general intelligence is not even general, but rather applies largely, although not exclusively, to academic kinds of tasks.” (Sternberg 2000b)

He is not without his detractors. In particular, Gottfredson (2002a, b), in a *magnum opus* of 54 pages, with 102 references, has systematically disputed Sternberg's central claims. The criticism is twofold: (a) practical intelligence is not distinct from traditional general intelligence, and (b) Sternberg offers insufficient evidence that practical intelligence is any better predictor of performance outcomes than general intelligence. Along the way, she shows evidence that he systematically distorts his own data, and frequently contradicts his own findings. The abstract says it all:

“The evidence [for practical intelligence] collapses however, upon close examination. Their key theoretical propositions are made plausible only by ignoring the considerable evidence contradicting them. Their six key empirical claims rest primarily on the illusion of evidence, which is enhanced by selective reporting of results. Their small set of usually poorly documented students cannot... dethrone g as the only highly general mental ability or intelligence.” (2002a, p. 343)

Why, then, is the idea of multiple intelligence so attractive? It appears that, somewhat counter intuitively, we academics are suspicious of “book learning”. Further, we are often loath to admit of stable differences in intelligence between people. But is there any truth to the idea that “g”, and IQ generally, are only useful measures of analytical skills or academic performance.

The notorious book, “The Bell Curve” (Herrnstein and Murray 1993), provided lots of data that showed that IQ predicted a large range of “practical” outcomes. Of course, there is a strong relationship between IQ and academic outcomes such as educational attainment. But IQ was also shown to relate to such diverse outcomes as health-related occupational disability (p. 161) unemployment, disability, illegitimate births (2% for the top quintile vs. 32% for the bottom) and criminality. Perhaps regrettably, the book then degenerated into a discussion of racial differences and an obvious right-wing agenda emerged. However, studies of predictive validity of IQ are not restricted to right wing journalists (Herrnstein, by the way, was, like Sternberg, a highly regarded psychologist). In fact, review articles regularly appear showing that IQ tests and aptitude tests are relatively strong predictors of academic and job success (Barrett and Depinet 1991; Sackett et al. 2008). As well as examining the predictive validity of aptitude tests, these authors also address issues of racial, social class and cultural bias, showing that tests, whatever their strengths and weaknesses, are not culturally biased. A complete review of these reviews is for another time, but in passing I was amused to find that, while driving tests do not predict accidents

or violations, paper and pencil tests do. Authentic and performance assessment are not necessarily superior predictors. Perhaps because doing well at parallel parking, the unanimously most difficult part of the driving test, is not a particularly good predictor of real driving skill. After all, few people are killed each year from poor parallel parking.

One other form of intelligence that is of particular interest to Lane is Emotional Intelligence, EI. Again, in the health care environment, a concern with the interpersonal dimension is natural. As she shows, communication skills influence compliance (on the positive side) and malpractice suits (on the negative side). So if we can find paper and pencil tests to measure people's ability to relate to others, this would fill a major gap in our admissions procedures.

But is EI that tool? Lane does cite one study showing that nurse's assessment of physicians' EI was a better predictor of patients' trust than the doctor's self-assessment (Weng et al. 2008). But that's part of the rub. EI is designed as a self-assessment exercise, and these data, among others, show poor predictive validity of the self-judged EI. As to the nurses, their assessment was likely a global rating of interpersonal skills, which may well be related to trust. But interpersonal skills are not synonymous with intelligence, emotional or otherwise, regardless of what the rating scale called it. To qualify as an intelligence, there must be evidence of stability since the first decade of life; not something assessed on the spot. Several reviews have highlighted the loose definition of EI, which on the one hand, appears to be trainable, so not a form of intelligence, and on the other, appears to be strongly related to personality, which then suggests it is not a measure of aptitude (Lewis et al. 2005; Romanelli et al. 2006; Davies et al. 1998). EI, after all is said and done, appears unrelated to career choice (Borges et al. 2009), academic performance (Romanelli et al. 2006; Newsome et al. 2000) or doctor—patient relationship (Weng et al. 2008). It remains to be shown if it has any value in health sciences education.

Personality

Finally, Lane ventures into the domain of personality. Again, she is, I think, severely handicapped by an underspecified theory of non-cognitive skill or aptitude. Personality psychology has moved toward acceptance of a five-factor model of personality, which includes Neuroticism, Extroversion, Openness to Experience, Agreeableness, and Conscientiousness. A number of measures of these traits exist, which, unlike measures of EI, appear well-validated. But when one moves to predictive validity, it appears that the most stable relationship is between conscientiousness and academic success, although the correlation with grades in medical or dental school are typically lower than 0.25 (Kulasegaram et al. 2009). Lane cites other studies showing a relationship to grades and academic success. But the whole point of the exercise is to find measures that predict *non-* academic performance. A demonstration that conscientiousness, or any other personality variable, predicts academic success is both theoretically barren, and practically useless, since previous grades, which are available “for free,” are better predictors of academic success—success begets success. Further, there is a regrettable tendency to correlate a more or less random set of personality variables with outcomes.. One study in her review nicely illustrates the point. The authors (van der Walt and Pickworth 2007) found that emotional stability, conscientiousness (yet again), self-discipline, and adaptability were positively correlated with academic success and anxiousness [sic], self-sufficiency, imaginativeness [sic] and openness were negatively correlated with academic success. The post-hoc explanation was that imagination might be counter-productive in some courses. But such

pos-hoc explanations are a better measure of the “imaginativeness” of the researchers than the participants.

Once we move beyond the relationship between conscientiousness and grades, the field becomes increasingly murky. Typically, a bunch of predictors including the ubiquitous “non-cognitive” measures are used to predict clinical ratings in clerkship or residency. A regression model is constructed where all of these are thrown into the pot and we see which sneak past $P < .05$. The assumption, presumably, is that clerkship ratings are more than just academic, although most studies of clerkship ratings suggest that the addition is just noise (Streiner 1985). Not surprisingly, given the methods employed, it is rare to find anything like the same variables recurring in different studies. From Lane’s review, we find that sociability, low life stress, low anxiety, internal locus of control, realistic direction, artistic direction, social direction, theoretical personality orientation (so much for “practical intelligence”), autonomy, complexity ... and on and on, are predictors of clerkship or residency ratings. Moreover, while some of these variables may sometimes become significant, close inspection frequently reveals small relationships. As one example, in the Keck et al. (1979) study cited by Lane, the regression model “included several non-cognitive measures”, but the individual correlations of the personality measures ranged from $-.24$ to $+.25$ and none were significant, and none accounted for more than 5% of the variance on the multiple regressions. By contrast, the correlation with GPA was 0.34 ($P < .006$) and with National Board scores was 0.39 ($P < .002$).

Such shotgun approaches will inevitably continue to yield confusing and unreplicable results. Perhaps the evolution of personality to the “Big 5” is a real sign of progress. But until we decide, one way or another, what we’re looking for, we will continue to create more heat than light. Do we want conscientiousness? Sounds good, but marks are a direct measure of this. Extraversion? But this is sometimes inversely related to conscientiousness, hence marks. And who wants a party animal for a doctor. Neuroticism—maybe a little. And so it goes. Unless we specify the desired personality traits we want in a doctor, hopefully by showing that they matter in terms of meaningful outcomes, we will continue to see one fishing expedition of correlations after another.

Conclusions

It is a truism to say that it takes more than academic grades to be a successful health professional. However, it is far easier to say that we need more qualities than it is to decide what these qualities are. And it is more difficult still to demonstrate empirically that these are both measurable and important.

The field has been plagued by a regrettable dualism, where everything is either a “cognitive skill” or a “non-cognitive skill”. There are two real problems with this duality. First, “non-cognitive” says what it isn’t (although even this is debatable) but not what it is. And when we get to what it might be, the breadth is so broad, ranging from motor skills, to non-analytic diagnostic reasoning, to interpersonal and communication skills, to personality, that it is unlikely that all can possibly have similar attributes. Second, just as we eventually abandoned the notion of “cognitive problem-solving skill” when we recognized content specificity, there is every reason to presume that such “non-cognitive skills” as communication or self-assessment will also prove to be highly content specific. Indeed, the success of admissions measures such as the Multiple Mini Interview (Eva et al. 2004), which explicitly acknowledge the limited generalizability of “skills” such as ethical

decision—making and communication, is empirical evidence against the existence of any kind of general and predictable non-cognitive skill.

Finally, as is evident from this critique, the field has been further hampered by dubious and wildly enthusiastic claims of proponents using buzzwords like emotional intelligence, learning style and personality that do not stand up to critical investigation.

There is much left to be done.

References

- Barrett, G. V., & Depinet, R. L. (1991). A reconsideration of testing for competence rather than intelligence. *American Psychologist*, *46*, 1012–1024.
- Borges, N. J., Stratton, T. D., & Elam, C. L. (2009). Emotional intelligence and medical specialty choice: Findings from three empirical studies. *Medical Education*, *43*, 565–572.
- Brooks, L. R. (1978). Decentralized control of categorization: The role of prior processing episodes. In U. Neisser (Ed.), *Concepts and conceptual development* (pp. 141–174). Cambridge: Cambridge University Press.
- Coffield, F., Moseley, D., Hall, E., & Ecclestone, K. (2004). *Learning styles and pedagogy in post-16 learning. A systematic and critical review*. London: Learning Skills Research Centre.
- Cook, D. A., Gelula, M. H., Dupras, D. M., & Schwartz, A. (2007). Instructional methods and cognitive and learning styles in web-based learning: report of two randomised trials. *Medical Education*, *41*, 897–905.
- Cook, D. A., Thompson, W. G., Thomas, K. G., & Thomas, M. R. (2009). Lack of interaction between sensing-intuitive learning styles and problem-first versus information-first instruction: A randomized crossover trial. *Advances in Health Sciences Education*, *14*, 79–90.
- Davies, M., Stankov, L., & Roberts, R. (1998). Emotional intelligence. In search of an elusive concept. *Journal of Personality and Social Psychology*, *75*, 989–1015.
- Eva, K. W. (2005). Dangerous personalities. *Advances in Health Sciences Education*, *10*, 275–277.
- Eva, K. W., Rosenfeld, J., Reiter, H. I., & Norman, G. R. (2004). An admissions OSCE: The multiple mini-interview. *Medical Education*, *38*, 314–326.
- Evans, J. St. B. (2008). Dual processing accounts of reasoning, judgment and social cognition. *Annual Reviews of Psychology*, *59*, 255–278.
- Gottfredson, L. S. (2002a). Dissecting practical intelligence theory: Its claims and evidence. *Intelligence*, *31*, 343–397.
- Gottfredson, L. S. (2002b). On Sternberg's reply to Gottfredson. *Intelligence*, *31*, 415–424.
- Herrnstein, R. J., & Murray, C. (1993). *The bell curve: Intelligence and class structure in American life*. New York: Free Press.
- Joughin, G. (1992). Cognitive style and adult learning principles. *International Journal of Lifelong Education*, *11*, 3–14.
- Keck, J. W., Arnold, L., Willoughby, L., & Calkins, V. (1979). Efficacy of cognitive/noncognitive measures in predicting resident-physician performance. *Journal of Medical Education*, *54*, 759–765.
- Kolb, D. A., Boyatzis, R. E., & Mainemelis, C. (1999). *Experiential learning theory: Previous research and new directions*. Working Paper 99-98. Cleveland, OH: Case Western Reserve University.
- Kulasegaram, M., Reiter, H., Hackett, R., Weisner, W., & Norman, G. (2009). Non-association between personality and non-cognitive skills in medical school admissions. *Advances in Health Sciences Education*, accepted.
- Lewis, N. J., Rees, C. E., Hudson, N., & Bleakley, A. (2005). Emotional Intelligence in medical education: measuring the unmeasurable. *Advances in Health Sciences Education*, *20*, 339–355.
- Newsome, Day & Catano, (2000). Assessing the predictive validity of emotional intelligence. *Personality and Individual Differences*, 1005–1016.
- Norman, G. R. (2008). When will learning style go out of style? *Advances in Health Sciences Education*, *14*, 1–4.
- Norman, G. R., Brooks, L. R., & Young, M. E. (2007). Non-analytical models of clinical reasoning: The role of experience. *Medical Education*, *41*, 1140–1145.
- Patel, V. L., Arocha, J. F., & Kaufman, D. R. (1999). Expertise and tacit knowledge in medicine. In R. Sternberg, J. A. Horvath, & D. R. Kaufman (Eds.), *Tacit knowledge in professional practice: Researcher and practitioner perspectives*. Mahwah, NJ: Lawrence Erlbaum.

- Reynolds, M. (1997). Learning styles: A critique. *Management Learning*, 28, 115–133.
- Romanelli, F., Cain, J., & Smith, K. M. (2006). Emotional intelligence as a predictor of academic and/or professional success. *American Journal of Pharmacy Education*, 70, 69.
- Sackett, P. R., Borneman, M. J., & Connelly, B. S. (2008). High stakes testing in higher education and employment. *American Psychologist*, 63, 215–227.
- Sternberg, R. J. (1984). What should intelligence tests test? Implications of a triarchic theory of intelligence for intelligence testing. *Educational Researcher*, 13, 5–15.
- Sternberg, R. J. (2000). Human intelligence: a case study of how more and more research can lead us to know less and less about a psychological phenomenon, until finally we know much less than we did before we started doing research. In E. Tulving (Ed.), *Memory, consciousness, and the brain: the Tallinn Conference* (pp. 363–373). Philadelphia, PA: Taylor and Francis, Psychology Group.
- Sternberg, R. J., Forsythe, G. B., Hedlund, J., Horvath, J. A., Wagner, R. K., Williams, W. M., et al. (2000). *Practical intelligence in everyday life*. New York: Cambridge University Press.
- Streiner, D. L. (1985). Global rating scales. In V. R. Neufeld & G. R. Norman (Eds.), *Assessing clinical competence*. New York: Springer.
- van der Walt, H. S., & Pickworth, G. (2007). Personality and academic performance of three cohorts of veterinary students in South Africa. *Journal of Veterinary Medical Education*, 34(3), 356–365.
- Weng, H. C., Chen, H. C., Chen, H. J., Lu, K., & Hung, S. Y. (2008). Doctors' emotional intelligence and the patient–doctor relationship. *Medical Education*, 42(7), 703–711.