Assessment in the interpersonal domain: Experiences from empathy assessment in medical education

Neville Chiavaroli
Australian Council for Educational Research

Neville Chiavaroli is a Principal Research Fellow in the Tertiary Education Program at the Australian Council for Educational Research, where his current research focuses on assessment and selection theory and practice in tertiary contexts. Prior to joining ACER in 2019, Neville was Head of Assessment in the Doctor of Medicine degree at the Melbourne Medical School, but his pedagogical work and research interests also encompass the development and assessment of empathy, professional identity formation and the medical humanities. He coordinated and taught the assessment subjects in the University of Melbourne's Graduate Certificate and Diploma courses for health professional educators for several years. Neville is an ongoing assessment consultant to many medical and health professional colleges and schools, and regularly conducts assessment workshops both nationally and internationally. Originally trained as a physiotherapist before completing a Master of Education and moving into educational research, Neville is currently pursuing a doctorate on learning and assessing in the affective domain in professional education.

Abstract

Frameworks for the teaching and assessment of 21st-century skills commonly recognise the importance of learning and skill development in the interpersonal domain. They also usually acknowledge the challenge of reliably and validly assessing students in this domain. In the field of medical education and in selecting students for medical courses, the concept of empathy has become central to representing the particular interpersonal understandings and skills expected of students and practising doctors. Attempts to assess these attributes during medical training are just as challenging as in school contexts. This presentation draws on several years' experience of working with medical educators to consider how empathy has been conceptualised, taught and assessed by educators. This analysis explores three common assessment approaches: self-report, performance examinations, and longitudinal observation and judgement in the clinical context. Each approach addresses important aspects of empathy and interpersonal skills. Each also has its limitations, although the self-report approach has emerged as the more widely known and used in medical education. Much still remains to be understood about making meaningful and valid use of observational judgements in the assessment of empathy, and, by extension, the interpersonal domain. In the meantime, useful guidance for teachers assessing interpersonal skills in the classroom may be found in alternative learning frameworks currently used in professional education that precede the 21st-century skills movement.

The interpersonal domain as a 21st-century skill

In 1970 the top three skills required by the Fortune 500 were the three Rs: reading, writing, and arithmetic. In 1999 the top three skills in demand were teamwork, problem-solving, and interpersonal skills. We need schools that are developing these skills.

Linda Darling-Hammond, Professor of Education, Stanford Graduate School of Education

The 21st-century-skills movement attempts to identify and promote the key skills that will support young people to successfully apply their learning to the world beyond their schooling. Alongside well-known skills such as critical thinking, problem-solving and personal motivation, frameworks for the teaching and assessment of 21st-century skills commonly recognise the importance of the interpersonal domain. The importance of such skills in life and in work seems undeniable, although their inclusion as a key skill for school curricula has been labelled as 'contentious' (Lamb, Maire, & Doecke, 2017). This paper will consider the approaches and implications for assessing this domain, based on the author's experience of working in medical education, where the promotion and monitoring of empathy is a key objective of medical courses.

The first thing to note is the diversity of terms used for skills in the interpersonal domain. A glance at the key 21st-century skills frameworks demonstrates the following terms being used by different educational reports: affectivity, collaboration, cooperation, (complex) communication, emotional learning, empathy, interpersonal domain/skills, relating to others, teamwork, as well as several variations on 'social' such as social awareness, social capability, social management, and so on. In medical education, these ideas are also referenced by concepts such as empathy, emotional intelligence, people skills, rapport, or 'soft skills'. This proliferation of terms can be confusing and frustrating, but they probably also point to the importance of the domain.

While a single, universally accepted definition of this construct or 'skillset' seems hard to come by, a succinct description offered by one educational body seems adequate and useful: 'skill in processing and interpreting both verbal and non-verbal information from others in order to respond appropriately' (NRC, 2011). The key terms in this definition are 'interpreting' and 'appropriately'. Good interpersonal skills involve insight, understanding, and the kind of situational awareness that helps one determine what might be an 'appropriate' response. There can be no set rules for determining this, much to the frustration of many – teachers and students alike. In other words, skill in the interpersonal domain involves some element of cognitive ability, a point explicitly made by Howard Gardner's (1983) coining of the term 'interpersonal intelligence'. Further, while it can be tempting to believe that people either have or do not have good interpersonal skills, 21st-century skill frameworks do not see it this way. As Lamb et al. (2017) succinctly note, two key principles underlie the conception of skills in frameworks: as 'developing expertise', and as 'contextual'. Both principles apply to the way the interpersonal domain is conceptualised and, necessarily, assessed (Spitzberg, 2003).

When it comes to the assessment of interpersonal skills, most 21st-century frameworks readily acknowledge the challenge this domain presents. Besides the elusive terminology, the frameworks also note the difficulty of precise assessment for such a 'complex' domain, the strong influence of context (including cultural), and the evolving nature of interpersonal skills in an increasingly sophisticated technological world (NRC, 2011). To some extent, these challenges apply to all the 21st-century skills, but particularly those 'complex skillsets', such as collaboration, which draws on multiple domains, including the cognitive and the social (Care & Kim, 2018).

However, it is worth remembering that educators have been wrestling with teaching and assessing in the domain long before the 21st-century-skills movement, and that useful guidance may be found in learning frameworks and taxonomies that have long been used in school contexts, and occasionally in professional contexts, too. The most obvious is Bloom's Taxonomy of the Cognitive Domain (Bloom, Engelhart, Furst, Hill, & Krathwohl, 1956; Anderson & Krathwohl, 2001), which outlines the different levels at which educational objectives can be focused and assessed with suitably adapted formats. Most teachers will be familiar with this framework, and it can be readily applied to the cognitive dimension of interpersonal skills. Less well-known is Krathwohl's Taxonomy of the Affective Domain, which provides a similar structuring for 'objectives which emphasize a feeling tone, an emotion, or a degree of acceptance or rejection' (Krathwohl, Bloom and Masia, 1964). A more recent taxonomy of interpersonal skills is that of Klein, DeRouin, & Salas (2006), which divides this domain into two broad areas, with associated subskills, as shown in Table 1.

Table 1 Taxonomy of interpersonal skills (Source: Klein et al., 2006)

Communication skills	Relationship-building skills
Active listening	Cooperation and coordination
Oral communication	Trust
Written communication	Intercultural sensitivity
Assertive communication	Service orientation
Non-verbal communication	Self-preservation
	Social influence
	Conflict resolution and negotiation

The example of empathy in medicine

> The biggest deficit that we have in our society and in the world right now is an empathy deficit. We are in great need of people being able to stand in somebody else's shoes and see the world through their eyes.

Barack Obama, 44th President of the United States

A related approach may be seen in the area of medical education, where the assessment of empathy represents a strong valuing of the interpersonal domain. In many ways, empathy is an ideal example with which to examine teaching and assessing in the interpersonal domain more closely. It is commonly acknowledged as involving multiple dimensions, for example, a *cognitive* dimension, which enables a person to *understand* the feelings or viewpoint of another, and an *affective* one, which allows a person to *feel* and *respond* to what the other may be feeling (Jeffrey, 2016); thus empathy would be classed as a 'complex skillset' (Care & Kim, 2018) in 21st-century frameworks. Similar to the status of interpersonal skills in these frameworks, empathy resonates strongly with stakeholders in medical education. For many, the concept of empathy has come to represent the particular interpersonal understandings and skills expected of students and practising doctors. In some cases, its deficiency is identified as a fundamental source of medicine's failures, as in the Stafford Hospital scandal of 2008 (Francis, 2013); or, indeed, society's failures, as the above quote by Barack Obama suggests. Assessing empathy in students, validly and authentically, is therefore vital.

Assessment approaches

Broadly speaking, there are three approaches to assessing empathy in medicine – self-report, direct observation (usually under examination conditions), and clinical supervisor judgement (usually longitudinal observation). The observation methods are sometimes referred to as 'third person assessments' (complementing the 'first person' perspective of the self-report measures) (Hemmerdinger, Stoddart, & Lilford, 2007); this highlights another possible approach to its assessment, termed 'second person', that is, the person who is on the receiving side of the interaction. In medicine, this 'other person' is usually the patient or their family, who, perhaps surprisingly, is only occasionally consulted as a source of judgement regarding students' (or clinicians') level of empathy. These potential approaches combine with three key considerations about assessing skills to determine how empathy is assessed in the clinical education context: ways to conceptualise a skill set, its contextual nature and the importance of authenticity of assessment.

Conceptualising empathy

There is a fundamental distinction between empathy as a form of understanding and as a form of feeling; in medicine, there is also an important third aspect – that of empathy-related *action*. This third dimension is often referred to as behavioural or communicative empathy. In other words, in medicine empathy entails thinking, feeling and behaving (Jeffrey, 2016). Sometimes a fourth dimension is defined: the ethical or moral dimension, specific to the role that empathy plays in compassionate care (Jeffrey, 2016). Clearly, empathy constitutes exactly the kind of 'complex skill set' discussed in 21st-century frameworks (Care & Kim, 2018).

Different emphases (or omissions) in relation to these three domains will affect the way empathy is assessed, or rather, the validity of any conclusion drawn from those assessments (Downing, 2003). This is an important issue in medical education. A recent review of empathy assessment in medical education (Sulzer, Feinstein, & Wendland, 2016) identified significant variation in the way different assessment methods defined or characterised empathy, along the three lines indicated above. Table 2 shows the relative emphases of studies that used available empathy measures for assessment purposes.

While the emphasis reflected in Table 2 is consistent with the place of empathy in medical education – most commonly understanding the patient's perspective, with acknowledgement that this understanding should lead to appropriate action by the doctor – Sulzer et al. (2016) noted that the selection of assessment instrument did not always match the dimension of empathy they were interested in. Clearly, there needs to be alignment between the underlying conceptualisation, as reflected in the objectives, and assessment methods for valid inferences to be drawn about student empathy development.

Table 2 Characterisations of empathy in available measures (Source: Based on Sulzer et al., 2016)

Empathy characterised as …	Studies (no.)
Thinking and acting	31
Thinking only	17
Thinking and feeling	14
Thinking, acting and feeling	12
Acting only	9
Acting and feeling	3
Feeling only	3

Contextual basis

Empathy, like the interpersonal skills domain, is generally acknowledged to be a contextual skill (Jeffrey, 2016), so that the nature and quality of empathy displayed by students depends on the given circumstances. Quality, in the interpersonal domain, is best summed up as 'effectiveness' and 'appropriateness' (Spitzberg, 2003); and the same author helpfully delineates the common contextual factors as culture, time (arguably 'timing' would be the better term), relationship, situation and function. Medical students learning the art and skill of empathy are often caught out by such contextual nuances; where the common phrase 'that must be really hard for you' might in some circumstances convey authentic empathy to a patient narrating her experience of illness, its over-use or hasty use, however well-intentioned, at the wrong time, or with the wrong patient, can have exactly the opposite effect (Coulehan et al., 2001). These factors impact on how empathy will be assessed, and judged, especially in the often summative and high-stakes context of medical school. Rubrics can be designed to support and guide assessor judgement on any particular assessment (Jonsson and Svingby, 2007), but they risk over-prescribing acceptable performance of such a complex skill.

Medical education's answer to this dilemma has been twofold: first, to assess empathy (along with other clinical skills) partly under standardised conditions with a highly-structured assessment format using trained, simulated patients, known universally throughout medicine as the OSCE (Objective Structured Clinical Assessment). Second, to draw on the key principle of sampling (Norman, 2002); that is, to assess empathy often, with different patients, in different clinical contexts, and by different assessors, thereby minimising the context-specific effects of the individual assessments. As one assessment expert puts it, referring to the measurement error inevitably contained in highly specific, contextual and necessarily unstandardised individual assessments, 'many fallible judgements, summed together, create value' (Hodges, 2013). While the notion of broad sampling would seem readily transferable to classroom contexts, the creation of a discipline-wide method of assessment of interpersonal skills would, I imagine, be prohibitive. Fortunately, it is neither desirable nor necessary.

Degree of authenticity

The significant advantage of the sampling approach is that it meets the third fundamental element of empathy assessment in medicine, namely authenticity. This notion is fundamental to the assessment of all 21st-century skills (Care & Kim, 2018), and in a practically-oriented profession such as medicine, is a key consideration in the evaluation of such skills, including empathy. In medicine, the strongest and most influential articulation of the goal of authenticity in assessment is represented by the taxonomy known as Miller's Pyramid (Miller, 1990)

This framework for assessment depicts visually the different 'levels' of clinical knowledge and skills desirable in medicine: *knowing*, *knowing how*, *showing how*, and *doing* – usually accompanied by common assessment methods targeting that level (see an example in Figure 1). In many ways this relatively simple framework is a variant of Bloom's Taxonomy, and no doubt other similar heuristics for teaching and assessment exist in classrooms both in Australia and around the world. But its impact in medical education has been profound, and has been credited with moving the practice of assessment from a poorly considered dependence on multiple choice questions and essays, to a more thoughtful alignment of assessment purpose, desired skill set and appropriate format. In other words, improving the authenticity, and potential validity, of assessments in medical education.
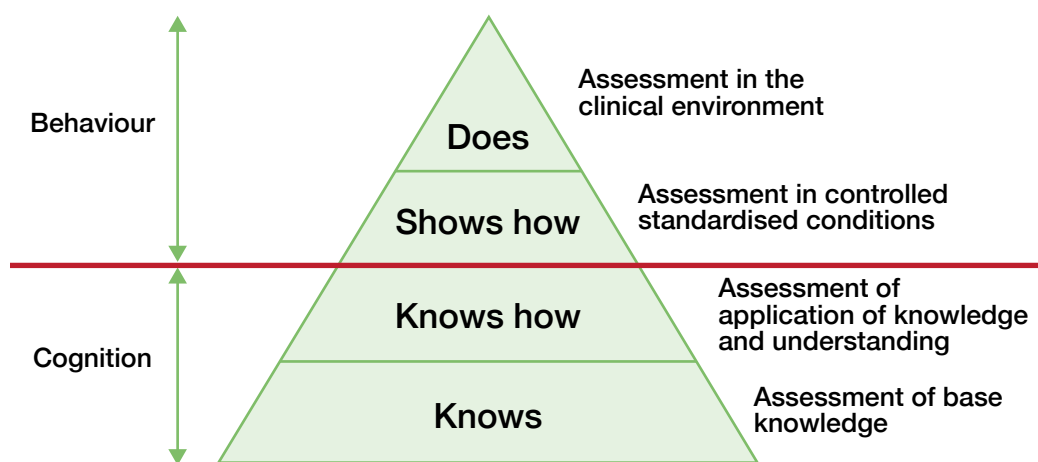

Figure 1 Miller's Pyramid of clinical assessment (Source: Adapted from Miller, 1990)

Applying this model to the assessment of empathy helps us make sense of the various conceptualisations, assessment approaches and tools available for assessing empathy and other interpersonal constructs. The cognitive dimension of empathy, understanding how others may feel or why they behave in a certain way in a given situation, can be represented by the levels of *knowing* and *knowing how*. The *knows* level aligns with an interest in students' base knowledge of human behaviour, assessed, for example through a written test, or self-report questionnaire relating to the value of certain principles for clinical practice.

The *knows how* level enables a higher level of contextual understanding and insight about people's thoughts and feelings. It can be assessed in written or oral formats, but clearly requires a specific context in which that understanding needs to be displayed. Commonly available commercial tests of empathy and related constructs such as the Mayer–Salovey–Caruso Emotional Intelligence Test™ and Ickes' empathic accuracy test target the *knows how* level are, but similar items, either selected or constructed response, can also be developed for classroom or clinical placement use.

At the top two levels of Miller's Pyramid, empathy is assessed as an action or behaviour, though founded upon the 'lower level' knowledge and understanding. *Showing how* requires the demonstration of relevant empathy but in a relatively controlled and standardised setting, usually represented in medicine by the OSCE assessment format. However some self-report instruments and 'situational judgement tests' (e.g. Lievens, 2013) that invite respondents to indicate how they might respond in a given situation could also be described at assessing at this level. However, as discussed above, empathy cannot be limited to constrained and prescribed situations. For the assessment of empathy in more authentic contexts, students are assessed in their everyday interactions with real patients, during actual clinical interviews or procedures, normally assessed by their supervisor or other clinical staff, using previously validated rating forms. Such assessments are commonly 'opportunistic', although may be planned in advance. The distinguishing feature of assessment at this level of 'doing' is the authentic context, the unstructured environment, and once again, the opportunity for multiple samples of the behaviour of interest.

Notably, the affective dimension of empathy is not clearly represented in Miller's Pyramid. This is consistent with the assessment approach in medicine which tends to avoid direct exploration of the affective or emotional aspect of medical training. Many medical educators claim this is a 'blind spot' in medicine's approach to empathy (e.g. Halpern, 2001). As mentioned previously, Krathwohl's Taxonomy of the Affective Domain provides a model for which the development of affective empathy could be charted and assessed. Self-report instruments would constitute the 'base level' of the domain, involving awareness and receptivity to others' emotions.

An important lesson from the above schema of empathy assessment is that educators need to resist the temptation to simply reach for the most common or convenient assessment format available. Various 'empathy assessments' conceptualise empathy differently, and target different dimensions and levels. A mismatch in these factors will undermine validity and risk drawing inappropriate conclusions about students' empathy. Many in medical education argue this may well be behind the contentious claim that medical students appear to 'lose' empathy through their course – a judgement usually based on the administration of self-report instruments rather than actual performance and judgement in authentic situations (Colliver, 2010).

Like most disciplines, there can be a gap between theoretical assessment approaches and actual practice. While medical courses may not always meet the goals of the curriculum designers, their attempts to enact authentic, aligned and valid assessment of empathy can provide a useful example for school classrooms faced with the challenge of assessing the interpersonal skills of students. Despite the obvious contextual differences, the assessment of empathy in medical schools provides an important example of how an interpersonal skill is highly valued, and how existing frameworks can assist teachers to assess them.

References

Anderson, L.W., & Krathwohl, D. R. (Eds.). (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. New York, NY: Longman.

Bloom, B. S., Engelhart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. (1956). *Taxonomy of educational objectives: Handbook 1: Cognitive domain*. New York, NY: David McKay.

Care, E., & Kim, H. (2018). Assessment of twenty-first century skills: The issue of authenticity. In E. Care, P. E. Griffin, & M. R. Wilson (Eds.), *Assessment and teaching of 21st century skills : Research and applications*. (pp. 21–39). Cham, Switzerland: Springer International.

Colliver, J. A., Conlee, M. J., Verhulst, S. J., & Dorsey, J. K. (2010). Reports of the decline of empathy during medical education are greatly exaggerated: A reexamination of the research. *Academic Medicine*, *85*(4), 588–593.

Coulehan, J. L., Platt, F. W., Egener, B., Frankel, R., Lin, C. T., Lown, B., & Salazar, W. H. (2001). 'Let me see if i have this right ...': Words that help build empathy. *Annals of Internal Medicine*, *135*(3), 221–227.

Downing, S. M. (2003). Validity: On meaningful interpretation of assessment data. *Medical Education*, *37*(9), 830–7.

Francis, R. (2013). Report of the Mid Staffordshire NHS Foundation Trust Public Inquiry: Executive summary. London, UK: The Stationery Office.

Gardner, H. (1983). *Frames of mind: The theory of multiple intelligences*. New York, NY: Basic Books.

Halpern, J. (2001). *From detached concern to empathy: Humanizing medical practice*. New York, NY: Oxford University Press.

Hemmerdinger, J. M., Stoddart, S. D., & Lilford, R. J. (2007). A systematic review of tests of empathy in medicine. *BMC Medical Education*, *7*, 24.

Hodges, B. (2013). Assessment in the post-psychometric era: Learning to love the subjective and collective. *Medical Teacher, 35*(7), 564–568.

Jeffrey, D. (2016). Clarifying empathy: The first step to more humane clinical care. *British Journal of General Practice, 66*(643), e143–e145.

Jonsson, A., & Svingby, G. (2007). The use of scoring rubrics: Reliability, validity and educational consequences. *Educational Research Review*, *2*(2), 130–144.

Klein, C., DeRouin, R. E., & Salas, E. (2006). Uncovering workplace interpersonal skills: A review, framework, and research agenda. In G. P. Hodgkinson & J.K. Ford (Eds.), *International Review of Industrial and Organizational Psychology* (pp. 80–126). New York, NY: Wiley and Sons.

Krathwohl, D. R., Bloom, B. S., & Masia, B. B. (1964). *Taxonomy of educational objectives: Handbook II: The affective domain*. New York, NY: David McKay.

Lamb, S., Maire, Q., & Doecke, E. (2017). *Key skills for the 21st century: An evidence-based review. Project Report*. Sydney, NSW: NSW Department of Education.

Lievens, F. (2013). Adjusting medical school admission: Assessing interpersonal skills using situational judgement tests. *Medical Education*, *47*(2), 182–189.

Miller, G. E. (1990). The assessment of clinical skills/competence/performance. *Academic Medicine*, *65*, (9 Suppl), S63–S67.

National Research Council (NRC). (2011). *Assessing 21st century skills: Summary of a workshop*. Washington, DC: The National Academies Press.

Norman, G. (2002). The long case versus objective structured clinical examinations. *British Medical Journal*, *324*(7340), 748–749.

Spitzberg, B. H. (2003). Methods of interpersonal skills assessment. In J. O. Greene & B. R. Burleson (Eds.), *Handbook of communication and social interaction skills* (pp. 93–134). Mahwah, NJ: Lawrence Erlbaum Associates.

Sulzer, S. H., Feinstein, N. W., & Wendland, C. L. (2016). Assessing empathy development in medical education: A systematic review. *Medical Education*, *50*(3), 300–310.