

Towards effective evaluation and reform in medical education: a cognitive and learning sciences perspective

Vimla L. Patel · Nicole A. Yoskowitz · Jose F. Arocha

Received: 29 August 2007 / Accepted: 8 November 2007 / Published online: 24 January 2008
© Springer Science+Business Media B.V. 2007

Abstract Health professions education is dealing with major transformations in light of the changing nature of the health care delivery system, including the use of technology for “just in time” delivery of care, evidence-based practice, personalized medical care and learning, as health professionals strive to integrate biomedical advances and clinical practice. This has forced the medical education community to reassess the current teaching and learning practices and more importantly, the evaluation of the medical education process. There have been recent advances in cognitive and learning sciences theories, some of which can inform medical educators about best teaching and learning practices and their impact on the evaluation process. An understanding of these theories provides a sound rationale for choosing specific instructional strategies and choosing evaluation measures that assess the curricular objectives. The review begins with an overview of evaluation and assessment in education, followed by an overview of major theories from the cognitive and learning sciences. Next, the role of cognitive and learning sciences theories in informing the process of medical education evaluation is discussed, including its impact on student learning, performance and professional competence, as well as recommendations for reform of medical curricula based on such theories. The paper continues with the elaboration of current trends in health sciences education, particularly medical education, and

V. L. Patel (✉)

Department of Basic Medical Sciences, The University of Arizona College of Medicine-Phoenix,
Arizona Biomedical Collaborative, 425 N Fifth Street, Phoenix, AZ 85004, USA
e-mail: vimla@asu.edu

V. L. Patel

Center for Decision Making and Cognition, Department of Biomedical Informatics, Ira Fulton School
of Engineering, Arizona State University, Tempe, AZ, USA

N. A. Yoskowitz

Laboratory of Decision Making and Cognition, Department of Biomedical Informatics, Columbia
University, New York, NY, USA

J. F. Arocha

Department of Health Studies and Gerontology, University of Waterloo, Waterloo, ON, Canada

available evidence for the impact on student learning and performance as well as areas where more research is needed.

Keywords Medical education · Evaluation · Assessment · Clinical competence · Cognitive science · Learning science

Introduction

There have been recent advances in the cognitive and learning sciences that lend themselves to the issues of learning, performance and evaluation in clinical domains. This paper begins with a brief overview of evaluation and assessment in general, in view to define and clarify some concepts and terminologies. This is followed by an overview of major theories from the cognitive and learning sciences. Next, the role of the cognitive and learning sciences theories in the selection of evaluation methodologies for medical education is discussed, including the impact on student learning, performance and professional competence.

Evaluation is the process of determining the value, significance, or worth of (e.g., a program) usually by careful appraisal and study. The primary goal is to determine the effectiveness of a program for achieving pre-set priorities and objectives, and identifying strengths and weaknesses and points for revision. There are multiple approaches to evaluation; that is, perspectives and organizing frameworks that guide the development and conduct of the evaluation. Some of the major evaluation approaches include systems analysis, behavioral objectives, decision-making, goal-free, art criticism, professional review, quasi-legal, and case study. Approaches differ in their target methods, models, and tools used.

Evaluation design can make use of either quantitative, qualitative, or mixed methods. Qualitative designs focus on description and explanation in evaluation, whereas quantitative designs focus on validation (Hawkes and Santhya 2002). One mixed-methods design is the extended-term mixed-method (ETMM), which is used to collect evidence on the effectiveness of educational programs, rather than solely using the results of randomized field trials, which is indicated by the What Works Clearinghouse's (WWC) standards (Chatterji 2005). ETMM uses both qualitative and quantitative methods, as appropriate, to evaluate programs from infancy throughout their lifespan to determine effectiveness and what works. Usually, descriptive research is conducted on the program in the earlier stages through program adoption and implementation in the particular environment. This is followed by experimental studies of the program at a later stage.

In addition, there are numerous evaluation models in the literature. These models usually outline what evaluators do (descriptive model) or what they should do (prescriptive model). Most models are largely theoretical and incorporate both formative and summative evaluation. Models are informed by a specific approach.

The general goals of education include the (1) acquisition of knowledge, skills, attitudes, and behaviors in a domain; and (2) self-directed learning, where learners take responsibility of their own learning. Learning occurs informally through everyday experiences, although to be competent and proficient in a domain, one needs formal education and training. Over the last three or so decades, perspectives on the effectiveness of learning have moved from a focus on repetition and practice to a focus on understanding and application of knowledge.

The goals of education should inform the strategies used in the educational process and in curriculum¹ design. These goals and strategies should be based on educational theory in an effort to bridge the gap between educational theory and practice, ultimately leading to better outcomes and performance (Kaufman 2003). As an example, self-directed learning can be a teaching strategy as well as a goal of the learner. As a teaching strategy, learning and instruction would be organized so that the learning tasks are mostly under the learner's control. As a goal, learners can strive to take responsibility of their own learning. Skills that learners should practice to achieve self-directed learning include asking questions, critically examining new information, identifying one's own gaps in knowledge and skills, and reflecting on one's own learning. Although self-directed learning is a goal of education, students' self-assessments of their skills and deficiencies, specifically in medical education, have largely been poor indicators of actual competence and performance (Eva et al. 2004). However, teachers can also contribute to effective learning by promoting self-efficacy in the learners by modeling, defining a clear picture of the desired outcome, providing the necessary basic knowledge and skills for the task, providing guidance and corrective feedback, and giving learners opportunities for reflecting on their learning. Along these lines, and based on the ideas of constructivism, the teacher's role is that of a guide and facilitator of learning, not merely a transmitter of knowledge. This instructional method is known as scaffolding. Initially, learning involves a significant degree of external support, through early environmental structuring, especially for beginning students. Increasingly, as competence is attained, there is a decrease in environmental support leading to opportunities for internalized self-regulation, through guided apprenticeship. In this way, as the direct external support fades away, learning occurs increasingly under the control of the learner.

In order to meet the above stated goals of education, students can engage in reflection on their learning experiences in two ways: "reflection in action," which occurs during the experience and involves actively engaging in the new experience by applying past experiences and reasoning to this new situation; and "reflection on action," which occurs after the experience and involves thinking about past experiences and how they might influence future experiences and practice (Kaufman 2003).

Today, the specific purposes of evaluation in education include documenting learning performance and achievement, making administrative decisions for education programs, and promoting faculty for reward of teaching excellence. In light of these purposes, what knowledge and skills should be evaluated and how should they be evaluated?

The main components of evaluation in education include an assessment of learners and learning, and teachers (relative to specified teaching standards) and teaching content. To accurately capture student achievement and predict performance, assessment of students and teachers needs to occur either within the actual performance context where the tasks are ecologically valid (Baker 2007), or in the laboratory where there is increased psychological validity but a certain lack of ecological validity (e.g., Norcini et al. 2002; Norman 2005a).

The evaluation and assessment methods discussed in this section apply to program evaluation and evaluation of education, in general. Before discussing evaluation in medical education, an understanding is needed of the theories from the cognitive and learning

¹ A curriculum is a set of intentions and expectations (Posner and Rudnitsky 1982) or a "formal and informal content and process by which learners gain knowledge and understanding, develop skills, and alter attitudes, appreciations, and values under the auspices of that school" (Doll 1996).

sciences that will provide the necessary rationale for informing medical education evaluation and reform.

Overview of cognitive and learning sciences theories

One of the goals of learning is the ability to transfer already acquired knowledge to new and unfamiliar problems and situations (Eva et al. 1998; Holyoak 2005; Patel et al. 1993). Earlier research, conducted according to a behaviorist perspective, did not take into account the cognitive and individual differences in learners, but emphasized repetition and practice. Current perspectives emphasize the understanding and application of knowledge as crucial to effective learning. Current learning theories include the following sections.

Adaptive character of thought (ACT-R) theory

Anderson's ACT-R theory asserts that complex cognitive processes, such as clinical reasoning and diagnostic decision-making, are a result of the interaction between two types of knowledge: procedural (how to do something) and declarative (facts) (Anderson 1983, 1993, 1996; Anderson and Schunn 2000). According to this theory, one needs to have a sufficient amount of both types of knowledge about a concept (working knowledge as well as knowledge of action) in order to solve problems about this concept. The ACT-R theory has certain implications for education and instruction. Anderson and Schunn (2000) advocate extensive practice in order to develop a high level of competence, arguing that the time spent practicing specific skills is the most important factor for developing lifetime competencies, a concept that is in keeping with the importance given to deliberate practice in the development of expertise (Ericsson 2004) and the power of practice using multiple problems or cases to aid in transfer to new situations (e.g., Gentner et al. 2003; Holyoak 2005). However, practice may not develop competence if the wrong knowledge is being emphasized and learned. Thus, ongoing feedback on students' learning is needed. As a result, computer-based cognitive tutors have been developed that monitor students' learning and provide immediate feedback on students' weaknesses, thus increasing the efficiency of the learning process. This theory has a number of implications for learning from errors, the nature of effective feedback and development of sustained competency for education and training.

Cognitive load theory (CLT)

Cognitive load theory (CLT) (Sweller et al. 1998) attempts to characterize and account for the role of memory and the complexity of learning materials in the learning process. The theory makes use of a number of hypotheses about the structure of human memory. First, it assumes, as it has been shown in memory research, that working memory (WM) is limited in terms of the amount of information it can hold. Second, and in contrast to working memory, it assumes no limits to long-term memory (LTM). Third, it also assumes that LTM is organized in the form of schemata, which are mental structures that serve to organize information in typical ways; are easily retrievable from memory; are often automatic, requiring no effort to use; and are used to interpret new, unfamiliar information. With these assumptions, CLT has been used to design instructional interventions that help

ease the learning process by preventing or limiting the learner's high memory load, which can result from either of two sources: The kind and amount of information presented to the learner as part of the instructional intervention (called 'extraneous' cognitive load) and the complexity of the information itself (called 'intrinsic' cognitive load), such as the number of idea units inherent in the information and the interaction among those units. This theory has implications for how information is organized in working and long-term memory for storage and retrieval of relevant information in a timely way. It also addresses the concerns of information overload during learning and during multitasking in clinical practice. More specifically, this theory is very important for the design of e-learning programs, as many extraneous components may be introduced that increase the cognitive load on the learner (Clark and Mayer 2007; Mayer and Moreno 2003). Clark, Mayer, Moreno and others have conducted several studies on the application of cognitive load theory to e-learning (e.g., Mayer et al. 2001; see also 2005 Special Issue in *Educational Technology Research and Development*, Vol. 53, Issue 3).

Situative learning theory

"Situative" learning theory focuses mainly on the performance of groups, rather than individuals (Greeno 2006; Lave 1988; Lave and Wenger 1991; Suchman 1985). One main principle of the situative approach is that learning is context-dependent (although not exclusively; see Anderson et al. 1996), which means that what is learned depends on the specific environment in which it is taught, and meaning is actively constructed within the specific learning environment. In other words, all interaction is actively constructed and negotiated by the learners using the available information and materials (termed artifacts) within the context of the learning activity. In addition, there is the opportunity for learning in any social organized activity, although this may not be formal and structured learning. In a study conducted from the situative perspective, analysis focuses on the performance of the whole activity system, which is a group of people and other systems, and individual cognition is only considered in relation to the group's patterns of interaction. "The goal is to understand cognition as the interaction among subjects and tools in the context of an activity" (Greeno 2006, p. 84). This has also been referred to as "distributed cognition." In the situative study, data are regarded and analyzed as records of interactions rather than verbal reports of one's thought and reasoning processes.

Based on the principles of situative theory, learning environments should be designed to promote learning of the desired knowledge and achievement of specific educational goals (Greeno 1998). In the biomedical domain, situative learning theory appears to be most useful in the characterization of learning in real-world practice settings, where the nature of these environments and resources available are constantly changing. In addition, given the emphasis on social components of learning, this theory provides a useful framework for understanding collaboration and teamwork in medical practice in regards to how clinicians construct their representations of collaborative clinical practice.

Cognitive flexibility theory (CFT)

Cognitive flexibility theory (CFT) (Spiro et al. 1991) accounts for the nature of learning in complex and ill-structured domains, such as medicine [because in medical problem solving, the initial states, the definite goal state, and the necessary constraints are not well

known (Feltovich et al. 1993)]. Cognitive flexibility speaks of the learner's ability to adapt to the content and the conditions of learning in domains where problems can be analyzed along several dimensions, requiring the learning to be flexible in response to different task demands (Spiro et al. 1991). CFT is based on the notion of "constructivism." Although there are various versions of this concept, the concept refers to the position that learners develop their understanding of the world by constructing models of reality in their minds. When given a text or a problem, the learner constructs its meaning by using the given information in conjunction with one's prior knowledge to come to an adequate understanding, or representation, of the text or problem. CFT de-emphasizes the retrieval of already formed static knowledge structures, and focuses on the need to use one's knowledge and various sources of information to create new understandings and new representations. Thus, CFT involves constructive processing, which requires the flexible use of prior knowledge along with the given information, which varies on a case-by-case basis.

Both the cognitive and situative programs of research have resulted in important knowledge about human learning that can, and should, inform the other when designing effective learning environments and instructional methods. The notion of learning in context is clearly one of the most important messages for education and instructional training. However, when training is situationally bound, and no provisions are made to emphasize the conditions of transfer, generalizability from one situation to another may be impeded. Proponents of problem-based learning (PBL) have claimed that learning in context facilitates retrieval of knowledge, and thus most learning should be context-bound, where biomedical knowledge is taught in relation to specific clinical problems to ensure the integration of biomedical and clinical knowledge. However, although biomedical knowledge is indeed integrated into clinical problems in PBL situations, this integration is often so context-dependent that its transfer to other situations is difficult (Bransford et al. 2000; Norman 2005b; Patel et al. 1993; Schmidt et al. 1990). These problems speak of the need to understand how physicians can acquire basic competencies in clinical practice through the apprenticeship process, but there is also an equally pressing need to understand how expert physicians acquire robust abstract conceptual models that have generalizability across contexts.

In this section, we have provided an overview of some important cognitive learning theories. In the next section, we focus on evaluation in medical education with respect to the goals of medical education and professional performance, while showing how principles from the cognitive and learning theories reviewed can be used to support these goals.

Evaluation in medical education

Goals of medical education

The overall goals of medical education are (1) the acquisition of the knowledge, skills, attitudes and values required to perform professional medical tasks competently and safely; and (2) the development and continuous refinement of the basic clinical skills that are required to provide competent care throughout a lifetime of professional work (AAMC 2004, 2005a).

Regarding the first stated goal, according to general education principles, assessment of performance should drive learning as well as the establishment of basic minimum levels of competency (standards). Cognitive theories of complex learning include the notion of

conceptual competence, defined as the potential to flexibly employ concepts in a range of contexts. A *theory of competence* would imply that there is a specific reference or expert standard indicating the content and form of knowledge in a given domain. Deviations from the standard may result from a lack of knowledge—as it is often assumed in traditional assessment—or from biases in reasoning or misconceptions (Feltovich et al. 1993; Patel et al. 1989).

In order to achieve this first goal of medical education, there needs to be a better integration between clinical experiences and the basic science courses. For example, early introduction of clinical experiences in the medical curriculum, commensurate with student experience and level of training, allows for the integration of basic science and clinical knowledge and for the formation and development of complex representations linking the basic science level to the clinical level and the intermediate levels of representation (Patel et al. 1991). This will enhance clinical reasoning if clinical symptoms are understood in relationship to underlying mechanisms. In addition, a “return to basic sciences” program in the last year of the medical curriculum would further help students consolidate their clinical knowledge with the necessary basic science knowledge. Such a program was instituted by McGill Medical School in 1977, called the “basic sciences options,” which was a 3-month program in the final year of medical school after the completion of the undergraduate clinical clerkship. The rationale for the program was that after 2 years of clinical experience, students would better appreciate the relevance of basic science information to their clinical practice, thus providing a chance for better integration of the two; as well as another opportunity for in-depth instruction in specific basic science areas (Patel and Dauphinee 1984). Patel and Dauphinee (1984) evaluated the program with respect to student learning and achievement and attitudes toward the program in 1978. The students were evaluated for their factual knowledge, integration of basic science with clinical knowledge, and interpretation of clinical data for decision-making; as well as a questionnaire to test student attitudes toward the program. The authors found that students held favorable views of the basic sciences program, and the major objectives of the program, which were to facilitate student learning in greater depth and to integrate basic sciences with clinical knowledge, were achieved. The structure of clinical problems learned through clinical practice provides a foundation to better integrate basic science information than vice versa (as reflected in all basic science learning before any clinical problem exposure).

Regarding the second stated goal, basic skill education and acquisition has been an implicit process, and it has been assumed that with repeated exposure to various situations, the appropriate clinical skills would be acquired. However, this is not necessarily the case. In addition, such clinical skills are not comprehensively evaluated. The AAMC Task Force on the Clinical Skill Education of Medical Students, established in 2003, defines “clinical method” as a “set of generic practice competencies required to provide medical care.” There are various, specific clinical skills which serve as the foundation for the generic competency. A clinical skill is defined as any “discrete and observable act of medical care” (AAMC 2005a). Skill learning requires the demonstration of skill proficiency. Assessment of skill proficiency requires observation to evaluate performance and learners need repeated and constructive feedback to improve and refine their skill proficiency (AAMC 2005a).

An issue dear to the cognitive perspective that is raised from the reported AAMC (2005a) view of skill learning is the assumption of generic competency that underlies both skill learning and skill assessment. As it has become evident from extensive cognitive research in learning and instruction, the existence of general skills and the possibility of

learning them have been questioned for several decades (Detterman and Sternberg 1993; Gick and Holyoak 1983; Salomon and Perkins 1989). These and other studies served to highlight the difficulty of learning general skills that can be applied in different contexts. Skills appear to depend not only on the context of practice—a position stressed by situative theory—but also on the specific underlying domain and task knowledge (see Ericsson et al. 2006). Although acknowledging these and similar criticisms of the concept of general skills, other researchers have pointed out that in some circumstances the learning of general skills (i.e., those that can be applied in dissimilar situations and problems) is possible if care is taken to carefully design situations that stress the critical aspects of similarity between situations and contexts. In this way, some degree of generality may be possible, if building on the concepts of reflective learning—which should ensure promotion of abstraction and self-reflection, as CFT proposes—and situative/distributed learning, making sure that the acquisition of effective skills takes place in the contexts and cultures of practice (Bereiter and Scardamalia 1996; Bransford and Schwarz 1999).

It is well known in learning sciences that students need to be exposed to various learning perspectives, including learning from clinicians in practice. Thus, teaching faculty should be diverse and include both generalists and specialists, and their teaching should be guided and determined by the educational objectives of the learning experience. Housestaff should be encouraged to teach and assess undergraduate medical students, thereby developing their own teaching and monitoring competencies as well as consolidating their own knowledge and skills, as is required in order to be able to teach them to someone else. These skills can only be developed through “learning by doing.”

How should we assess teachers and teaching to improve student performance? In order to answer this question, we need to define what makes a teacher effective. For example, some characteristics include being knowledgeable, skilled clinicians, curious, inquisitive life-long learners, excellent communicators and role models of professionalism, committed to developing their teaching skills, including the ability to assess the range of learners’ needs and to take the level of each learner into perspective (Weinberger and Whitcomb 2003). They also should be able to provide timely, constructive, and effective feedback to learners. This kind of feedback is known to change behavior more favorably than providing either unconstructive, delayed or no feedback.

Competence in medical education

Since the goal of medical education is to perform medical tasks competently, how do we define competence? *Professional competence* in medicine has been defined as the “acquisition of a strong and broad knowledge base, a range of clinical and professional skills, and exemplary professional and humanistic behaviors” (AAMC 2006). The goal of such competence is to have a positive impact on patient outcomes. Such competency could be achieved by learning through practice and reflection on experience (AAMC 2006).

What competencies are necessary and what knowledge, skills, and behaviors are required for competent performance? According to the AAMC, each medical education program needs to establish a set of graduation competencies and a specific set of skills for student achievement, which would then inform curriculum design to allow students to acquire these competencies (AAMC 2006). There are multiple competencies involved in medical performance, some of which are informally acquired in the context of practice, whereas others are best acquired through a formal learning process. Medical students may show competence in solving familiar problems because of well-organized and easily

accessible knowledge but may not show the same competence when dealing with unfamiliar or novel problems. This addresses the issue of viewing competence as an ability to be flexible and to transfer knowledge across problems and domains, as addressed in empirical studies in the cognitive and learning sciences. Competence can also be viewed as a function of level of training and amount of “deliberate practice” (Ericsson 2004) and reflection on one’s experience. Therefore, one would assume that seasoned physicians would have a higher level of competence than less experienced professionals, residents, or medical students. Conceptual competence develops through the deep understanding of general principles of a domain (Gelman and Greeno 1989), which is characterized by generativity—the ability to use knowledge in a variety of task and contexts—and robustness—the ability to adapt acquired concepts to unfamiliar task constraints or novel situations. The extent to which aspects of a domain are best learned in context is determined jointly by the nature of the domain knowledge and the kinds of tasks that are performed by practitioners. Such competence can be assessed by carefully designed assessment tools that capture the qualitative nature of students’ thought (thinking and reasoning) and action (providing problem solutions, e.g., diagnoses, and making decisions, e.g., patient management plans).

Fostering and assessing competence

Cognitive theories of complex learning shed light on how to assess competence by suggesting methods of testing that emphasize the flexibility inherent in conceptual understanding. It needs to be acknowledged that traditional methods of assessing achievement and competence are not sufficient for testing for flexible understanding with more difficult and complex material. Thus, instruction and assessment needs to be reformed to effectively test such deep understanding and flexible problem-solving. For example, medical instruction should include a diagnostic component to identify problems, where student’s preconceptions are identified and clarified, and a prescriptive component, where direct challenges to areas of knowledge that may present barriers to understanding are provided (Feltovich et al. 1993).

Another way to assess competence is to assess the reasoning strategies used when solving clinical problems. According to cognitive theory, the development of data-driven reasoning (as exemplified by experts) is associated with the development of automaticity of clinical skills (Patel et al. 2000b). To support this development, students need to acquire both biomedical and clinical knowledge. Clinical knowledge is generally sufficient to solve most routine problems in clinical practice. However, to solve difficult problems, biomedical knowledge is necessary, which is associated with the use of backward (or hypothesis-driven) reasoning, consisting of chains of causal explanations that serve to explain the biomedical rationale for the current difficult problem. The biomedical knowledge is fine-tuned to account for anomalous findings in the complex problem. It should be noted that biomedical knowledge can exist without any clinical context and it serves a specific purpose when provided in a clinical context. Thus, the type of reasoning and knowledge used when solving routine and difficult clinical problems can be an indicator of the level of competence (Patel et al. 2000a).

Assessment of competence in medicine has been typically based on the notions derived from the Bloom’s taxonomy of educational objectives (Bloom 1956). Using this taxonomy, evaluation of educational programs usually found objectives to be in the *Knowledge* category, thus emphasizing mere recognition or recall of information. However, the most

important educational objectives are generally considered to be related to the understanding and use of knowledge (i.e., cognitive processes, such as comprehension and synthesis). A revision of the taxonomy (Anderson and Krathwohl 2001; Krathwohl 2002) now provides a better conception of how educational goals and objectives can be linked to competencies and assessment of performance, by separating types of knowledge from the cognitive processes used. Using this revised taxonomy in the form of a table with Knowledge on the vertical axis and Cognitive Process on the horizontal axis, educational objectives can be classified according to the type of knowledge used as well as the cognitive process used to understand and use the knowledge. This table can be used in two important ways: (1) for evaluating instruction and teaching content by identifying which Knowledge \times Cognitive Process categories are lacking based on which categories are covered by the educational objectives; and (2) for classifying the learning activities used to achieve the outlined objectives and the assessments used to evaluate students' progress in achieving the objectives. This revised taxonomy has added ability to classify standards, in addition to educational goals and objectives.

In traditional forms of assessment, the presence of preconceptions and misconceptions in students' understanding is not typically a focus of attention. This type of assessment requires measures that tap understanding and application of clusters of related concepts, not just individual concepts (e.g., to test for flexibility), as is considered important in cognitive assessment. The cognitive flexibility perspective also requires de-emphasizing compartmentalization of knowledge, and focusing on the connection of multiple concepts and their interaction and variation across contexts, a goal that is inconsistent with the current view of hierarchical learning emphasized by behavioral perspectives on instruction and assessment, e.g., Bloom's taxonomy and cognitive theories of simple learning.

Assessment also needs to take into account the influence of contextual factors on performance (Epstein 2007). The rationale for the latter comes from assumptions of "situative" theory, which views the individual and environment as dynamically interacting, and focuses on learning and instruction that is context-based. However, as discussed earlier, an exclusive focus on learning in context may reduce future abilities to transfer knowledge to clinical problems in different contexts.

Teachers and teaching

It is well known that the quality of teaching influences student learning and performance, so assessment of teachers and teaching is critical to evaluation of an educational curriculum (e.g., Griffith et al. 1997, 1998; Jolly et al. 1996). For example, Patel and colleagues (1991) found that after the first year of medical school, students learning pulmonary physiology had difficulties in developing an accurate pathophysiological model of a given problem and were not able to correlate clinical symptoms with the underlying pathophysiological mechanisms. These difficulties were the result of the type of teaching (lecture format) and instruction, which did not make relationships between clinical symptoms and underlying mechanisms sufficiently explicit. Teachers need to impart robust basic science models to students so they can use these models for clinical problem solving. Although not discussed in this paper, the methods for evaluation of clinical teaching are similar to those used for evaluation of student performance (see Snell et al. 2000 for a review).

Student learning and performance

Assessment of learners is used to determine competence, predict performance (as student learning is closely tied to student performance), measure performance improvement after feedback, and give a summative grade (Metheny et al. 2005). A combination of multiple assessment methods is necessary to assess learning outcomes, and different methods are more useful for different learning outcomes. For example, Miller (1990; as described in Shumway and Harden 2003) proposed a pyramid of learning (from cognition at the bottom to behavior at the top) and van der Vleuten (1996) ordered assessment methods according to their appropriate place in the pyramid. In a similar way, Shumway and Harden (2003) placed their 12 identified learning outcomes within the pyramid. Combining both pyramids, the optimal assessment methods for each learning outcome are explicated. Taking this one step further, the learning outcomes and assessment methods could be mapped along the cognitive processes and knowledge dimensions of Bloom's revised taxonomy, as described earlier. In this way, type of knowledge learned, cognitive processes used, clinical performance and associated methods of assessment can be linked, providing a strong rationale for the educational experiences and instructional methods used. This would bring the focus of assessment to both knowledge acquired and its use in clinical performance, suggesting that knowledge-based assessments are not inferior to performance-based assessments, but that both are necessary in evaluating student learning. In a recent editorial, Norman (2005a) has suggested that acquired knowledge may be as, or even more, important in predicting actual physician performance, and as such, should not be regulated to the bottom of Miller's pyramid.

The method used to assess clinical competency has a strong influence on how and what students learn (Howley 2004). Therefore, methods of assessment are closely tied to the learning process and student performance, and ultimately patient outcomes. There is no assessment that will provide the perfect picture of an individual's knowledge, skills, and performance, but the assessment needs to be valid and reliable enough to provide useful data for evaluating clinical competence. Norman and Feightner (1981, p. 26) define clinical competence as "the ability to gather data from the patient by history and physical examination, integrate this information into a diagnostic formulation, select appropriate investigations to confirm the diagnosis, and institute efficacious management."

Howley (2004) describes three areas for the direction of performance assessments, including creating evidence-based locally developed assessments (valid in the context), an understanding of educational outcomes (influence and effectiveness of feedback) and non-cognitive assessment factors (student anxiety, motivation, learning styles), and more student-driven assessments [conducted in naturalistic environments (*authentic assessment*); transfer of knowledge and skills from classroom to bedside]. Recently, there have been studies investigating the effects of various factors on students' Objective Structured Clinical Examination (OSCE) performance. For example, Junger et al. (2005) found that a skills training course had a positive impact on OSCE performance. In addition, Blaszkiewicz et al. (2004) found that the testing context (obstetrics-gynecology or psychiatry rotation) influences OSCE performance (i.e., collection and interpretation of information regarding case), despite the fact that the OSCE station was identical in both contexts. This finding supports the need to use case scenarios that cover a broader range of diagnoses, not limited to that of the immediate context.

Feedback is another factor that has been shown to influence clinical performance. Veloski et al. (2006) conducted a systematic review of studies published between 1966 and April 2003 on this issue, asking the following questions: "What are the features and characteristics of feedback that influence physicians' clinical performance?" and "What

are the moderating factors of this relationship?” The authors showed that 70% of the studies reviewed (29 of 41) reported positive effects of feedback on physicians’ performance (where feedback was not combined with other interventions). The most important features of feedback that were positively associated with performance included the source of the feedback (administrative units and professional groups) and the duration of the follow-up period of monitoring physician performance and providing feedback (i.e., a duration of at least 2 years was associated with more positive outcomes than a duration of less than 1 year). These findings highlight the importance of timely and constructive feedback from teachers as well as long-term monitoring of student progress and performance. In addition, more studies need to be conducted to explore the predictive relationship between performance assessment during medical school and longer-term outcomes, as the current literature has only shown a low to moderate correlation between student assessment and postgraduate training performance (Hamdy et al. 2006).

Current trends in health sciences education

Four areas that are becoming prominent in health sciences education and medical care are (1) personalized medicine, (2) evidence-based practices, (3) inquiry-based learning, and (4) team-based learning. Personalized medicine aims to achieve optimal medical outcomes by helping physicians and patients choose the disease management approaches likely to work best in the context of a patient’s genetic and environmental profile and includes the use of new methods of molecular analysis to better manage a patient’s disease or predisposition toward a disease, genetic screening programs that more precisely diagnose diseases and their sub-types, or help physicians select the type and dose of medication best suited to a certain group of patients.

Evidence-based clinical practice is based on empirical findings of effective treatments and diagnoses and decision-making and is usually provided to physicians in the form of clinical practice guidelines (CPGs) that provide a means to manage “information overload” and support effective learning and performance. Unfortunately, CPGs are often underused. The literature on the use of CPGs cites ineffective dissemination, irrelevance of guidelines to the care of specific patients, and the substantive nature of the guideline information as factors most frequently related to the poor use of guidelines. Patel and colleagues (2001) investigated the effect of cognitive factors on the nature of experts’ and sub-experts’ (experts working outside their area of expertise) guideline use. Findings suggest that the format of guidelines may be differentially appropriate in different contexts: algorithmic guidelines are more appropriate for supporting real-time problem solving, and narrative guidelines are more appropriate for supporting learning. Clinical practice guidelines are frequently used as the knowledge base for computerized clinical decision support systems (CDSSs). A systematic review of the impact of CDSSs on practitioner performance found that better performance was documented in studies where the practitioner was automatically prompted to use the system than in studies where the practitioner had to actively initiate the system (Garg et al. 2005). In addition, clinical recommendations were only directly provided to the practitioner on the computer screen in 41% of the reviewed studies, although most studies did not assess the impact of the CDSS on clinician workflow.

Team work is increasingly recognized as central to effective patient care. However, physicians are not adequately trained to work in teams. Interprofessional education, the “learning together to promote collaborative practice” (Hammick 1998, as cited in Cooper

et al. 2001), is required. Cooper and colleagues (2001) reviewed the literature on educational interventions aimed at interprofessional education in undergraduate health profession programs. The studies identified that measured outcomes found mostly beneficial results. Specifically, interventions had the largest effect on students' knowledge, attitudes, skills and beliefs, especially on understanding professional roles and working in teams. Most outcomes were only assessed in the short-term, and there were few studies that summatively assessed student performance.

Relevant to the development of expert teams is the work of Guest and colleagues (2001), who discussed deliberate practice (Ericsson 2004) and sustained practice (quality use of time) to develop and maintain expertise (even in seasoned clinicians) in two types of tasks: static and dynamic. Deliberate practice on dynamic tasks should involve reflecting on the task, one's actions, and thought processes. Lack of motivation for continuing to improve one's skills and understanding halts the further development of expertise. Therefore, teachers need to encourage students to form a positive attitude toward work to address motivational issues down the line. However, deliberate practice appears to be more useful for relatively simple tasks, but not as useful for acquiring complex conceptual models.

Team performance

Medical education has historically focused on the individual physician's performance, but given the changes in the health care system, there is a move towards incorporating teamwork training, especially during residency, and assessing team performance. The situative approach is a useful basis for understanding team processes, communication and performance. As described earlier, the situated approach involves a shift from viewing cognition as a property of the individual to viewing cognition as a distributive property of individuals interacting with people and artifacts in the environment (Greeno 2006). In terms of medical education, integrating team-based learning in biomedical curricula becomes critical, especially in clinical situations where problem-solving requires cooperation and coordination among multiple team members, such as in the hospital environment (e.g., emergency and intensive care units). Team performance is further discussed in the section on simulation-based team training.

Cultural competence

Cultural competence has been defined as "a set of congruent behaviors, knowledge, attitudes, and policies that come together in a system, organization, or among professionals that enables effective work in cross-cultural situations" (AAMC 2005b, p. 1). The Liaison Committee on Medical Education (LCME) introduced the following standard for cultural competence in 2000:

The faculty and students must demonstrate an understanding of the manner in which people of diverse cultures and belief systems perceive health and illness and respond to various symptoms, diseases, and treatments. Medical students should learn to recognize and appropriately address gender and cultural biases in health care delivery, while considering first the health of the patient. (AAMC 2005b; p. 1)

In order to be effective and meet these standards, cultural competence education needs to be integrated into the existing medical school curriculum and assessment. For evaluation of students in cross-cultural education, the AAMC developed the Tool for Assessing Cultural Competence Training (TACCT), a self-administered assessment tool for medical school's use to examine all components of the medical school curriculum and aid medical schools in meeting the LCME objectives. The TACCT can be used in both conventional and problem-based curricula. The two parts to the TACCT include monitoring of where teaching is occurring, and what learning objectives are being met. However, the TACCT cannot be used for detailed analysis of teaching strategies or learning outcomes.

The University of Michigan Medical School implemented and evaluated an undergraduate Sociocultural Medicine Program (SMP) (one part of Sociocultural Medicine Curriculum), providing a successful and informative model on how to incorporate sociocultural medical training into medical education (Tang et al. 2002). Tang et al (2002) uses "sociocultural medicine" to refer to "the understanding, incorporation, and application of social and cultural issues in health, medicine, and patient care" (p. 578). This program is one of the few instituted in medical schools across the country. Changes in students' attitudes about sociocultural issues in medicine improved in the positive direction, although there was no significant change in students' perceptions of the impact of sociocultural background on specific clinical scenarios. This finding emphasizes that the knowledge of the influence of sociocultural issues in the clinical situations that students learned in the SMP program did not generalize to other clinical situations. Consistent with learning theories described earlier (i.e., Cognitive Flexibility Theory), Feltovich and colleagues' recommendations for the use of multiple, overlapping cases for teaching purposes is very useful.

Other studies related to cross-cultural education and cultural competency training include Betancourt (2003), Tervalon (2003), and Hobgood and colleagues (2006), who review educational models and evaluation methods of cross-cultural competency education in the health professions, and Gregg and Saha (2006), who discuss the uses and misuses of culture in medical education and warn against teaching and promoting a simplistic view of culture, which will enforce the use of racial and ethnic stereotypes and biases in medical care.

Simulation-based learning

There has been an increasing trend in medical education towards interactive learning. Simulation techniques are a great example of this new trend (see Magee 2006 for a review of simulation in adult education). The main benefits of this approach include, but are not limited to, having a clinical setting without risk of causing patient harm and having a learner-centered environment. Students learning through simulation have expressed a feeling of increased clinical competency as well as enthusiasm for the learning material. Gaba (2004) advocates the use of simulation for the improvement of patient safety and patient care using various applications. The motivation for the use of simulation for training and education in healthcare is the effective use of simulation in other industries, such as the military and commercial aviation. Gaba (2004) outlines several diverse uses of simulation in healthcare, including (1) education and training of clinicians, (2) assessment of clinicians' performance, (3) research and evaluation of organizational practices and exploration of human factors (e.g., fatigue), (4) the usability of clinical equipment, and (5) as a tool for helping create the desired "culture of safety" in the workplace. In addition, for learning purposes, simulation may be used to help students acquire new knowledge and

build deeper conceptual relations between discrete pieces of knowledge, as well as integrate learning basic concepts with performance and development of the appropriate clinical skills. One major benefit of using simulations is the training in a risk-free environment, which allows students to make errors and learn from them without the fear of hurting a patient (Friedrich 2002). This is especially important for crisis training and training for clinicians in emergency medicine (discussed in a later section). However, there needs to be evaluation of simulation-based learning and training that shows increases and efficiency in learning (Friedrich 2002) in addition to the nature of learning that takes place with the use of technology.

Comparison with other curricula and teaching methods

Simulation-based learning has been recently compared in empirical studies to other learning approaches, such as conventional and problem-based learning approaches. Steadman and colleagues (2006) evaluated and compared the performance of students using simulation (SIM) vs. students using problem based learning (PBL) techniques. Thirty-one fourth year medical students participated in this study via a 1 week acute care course. After having their critical care skills assessed, students were randomly assigned to either the PBL or SIM group to learn about dyspnea. To equilibrate the time spent for simulator use, the PBL group used the simulator to learn about acute abdominal pain, while the SIM groups utilized PBL techniques to learn about acute abdominal pain. Standardized checklists were used to assess student performance at the end of the week on a unique dyspnea scenario. Results indicated that the SIM group performed significantly better on this final assessment than the PBL group. One explanation for this result is that the simulator experience requires the learner to be more engaged in learning, activating multiple learning pathways (e.g., auditory, visual, and tactile).

In a randomized controlled trial of third and fourth year medical students, simulation-based teaching was compared to traditional teaching methods through the use of pre–post written evaluations (short-answer questions) on both domains of reactive airways disease and myocardial infarction (Gordon et al. 2006). In a single session, students received either a simulation on myocardial infarction followed by a lecture on reactive airways disease, or a simulation on reactive airways disease followed by a lecture on myocardial infarction. Although both methods of teaching resulted in improvement in scores on the short-answer questions examination for both domains, there were no significant differences between the groups in their performance on the examinations in either subject domain. However, this study was only conducted as a single teaching session and the long-term impact was not assessed.

Impact on student learning and performance

Simulations are increasingly used in medical education, teaching and assessment, although their use is not very widespread. An international survey (Morgan and Cleave-Hogg 2002) of the use of simulation in education, evaluation, and research in anesthesia found that most simulation centers responding to the survey (38%) used simulation in undergraduate and postgraduate teaching, but very few centers used simulation for evaluation or practice assessment. The biggest challenge to incorporating simulation into curricula is the high cost of equipment and their maintenance.

Despite their light utilization, simulations have been shown to positively impact learning. In a recent review of medical simulations (Issenberg et al. 2005), it was found that under certain conditions, high-fidelity medical simulations facilitate learning. The review outlines 10 ideal conditions for educational programs that maximize the effectiveness of simulation in education. These conditions include providing educational feedback, participating in repetitive practice, using curriculum integration, using tasks from a range of difficulty levels, employing multiple learning strategies, capturing clinical variation in simulations, using a controlled and safe environment for training, having individualized learning experiences where students are active learners, having clearly defined goals and outcomes, and using valid simulators. Simulation training does not and should not replace real patient contact, but instead complements and prepares learners for direct patient care and allows them to acquire skills in a controlled and safe environment, leading to increases in self-confidence which may affect clinical competency.

A recent pilot study (Gordon et al. 2006) explored if a brief exposure to patient simulation during a medical student's pre-clinical years is able to enhance their ability to learn basic concepts in cardiovascular physiology. Students belonged either to a control group (discussion only) or intervention group (discussion and simulation) of a standardized case of myocardial infarction. Students were then evaluated with a brief multiple choice test a few days after the case discussion (control group), immediately after simulation (intervention group) and 1 year later (both groups). Results indicated that the simulation group showed higher scores both immediately and 1 year after intervention.

Patient simulators have been tested for use in the evaluation of performance in such areas as anesthesiology. One study (Devitt et al. 1998), found that a performance rating scale was able to discriminate between resident and faculty anesthesiologists. The current challenge for the use of simulators is the ability to create scenarios that are valid and reliable in assessing performance on a variety of tasks, while capturing the complexity of doctor-patient interactions (Kapur and Steadman 1998). Furthermore, it is not quite clear what the nature of the learning is that is acquired through simulation. However, although much research needs to be conducted on the uses and effects of simulations on medical learning, it is possible to speculate about the reasons simulations may be more effective than traditional methods for fostering learning and improving performance. Going as far back as to the work of the early constructivists, such as Piaget and Vygotsky, a central idea has been that knowledge and skill acquisition through engaging actively in learning (e.g., through manipulation) opens up the possibility for reflection and better understanding. Given the dynamic nature of simulations, and therefore the opportunities for active engagement, they appear to afford the possibility of developing the needed cognitive flexibility to adapt performance to the changing tasks demands, as suggested by CFT as critical for complex learning (Spiro et al. 1991). It is important to note, however, that simulations in and of themselves may not be sufficient to foster learning, as the design of instruction would need to be guided by cognitive learning principles, such as those in CFT, to be most useful and effective, which requires deeper forms of assessment (e.g., conceptual understanding) than those used in traditional evaluations (e.g., rating scales, multiple choice items).

Simulation in surgical training

It is well known that simulations are good ways to acquire skills. Simulation-based training in surgery provides an enormous opportunity for residents to develop and refine

their surgical skills in a variety of scenarios, without risk to patients (Pellegrini 2006). Even more than in other medical specialties, the practice of surgery requires a high degree of precision and attention to detail. Simulators that have been developed for use in surgical training include mechanical simulators, computer-based simulators, virtual reality environments, and hybrid simulators, all of which benefit the acquisition of different types of skills. Several studies indicate that surgery skills acquired during simulation training (both low-fidelity and high-fidelity simulations) transfer to the operating room (e.g., Fried et al. 2004; Gallagher and Cates 2004; Seymour et al. 2002), and some studies have even shown that low-fidelity simulations transfer as well as high-fidelity simulations (Hamstra et al. 2006). However, although there are many studies that have shown the value of its use in surgery, this issue continues to be debated (see Sutherland et al. 2006 for a review and Dutta et al. 2006 for critical commentary).

Simulation in emergency medicine training

Simulation is being increasingly used in graduate medical education (McLaughlin et al. 2006) to ensure patient safety and as a measure of professional competence (see Norcini and McKinley 2007 for a comprehensive review of assessment methods). A recent national survey explored the status of simulation training in emergency medicine by collecting data about the type of simulations used from 126 residencies (Bond et al. 2007). Results indicated that the most common use of simulators is Advanced Cardiac Life Support (ACLS) mannequins, followed by high-fidelity mannequin-based (HFMB) simulators.

High-fidelity simulations have been used to evaluate students' and residents' development of the ACGME's domains of competency, such as the systems-based practice (SBP) competency. The SBP competency is defined as "an awareness of and responsiveness to the larger context and system of health care and the ability to effectively call on system resources to provide care that is of optimal value" (ACGME 1999). For emergency medicine physicians, having expertise in the systems-based practice domain is critical because they interact with many of the departments and services within the hospital as well as the health care system outside of the hospital and they evaluate and treat the entire spectrum of the patient population. Faculty at Northwestern University's Emergency Medicine Residency Program developed and implemented a simulation-based curriculum to address the systems-based practice competency, which can be used to enhance case-based learning and provide another tool for evaluation of residents' performance (Wang and Vozenilek 2005). Other simulation-based curricula have also been proposed (e.g., McLaughlin et al. 2002). With an increased focus on competency and outcomes in the curriculum, to what extent is knowledge that is tied to the skills that are being acquired also developed? It is not clear that assessment of performance in simulations and clinical scenarios will be adequate to understand the quality and depth of knowledge that is or is not being developed. This is where cognitive studies of clinical reasoning and problem-solving are essential for identifying students' progress in biomedical and clinical knowledge acquisition and organization in relationship to the development of clinical skills. Although several cognitive studies have investigated such issues, as new curricula are being introduced and changes are made in current programs, these studies are needed in order to adequately evaluate the effectiveness of the curricula reform and to identify areas for improvement.

Simulation-based team training

Simulators are often used in group settings, with many programs offering team-based simulation training (Baker et al. 2005). Research has shown that such training may significantly increase teamwork performance (Wallin et al. 2007); and that didactic teamwork training can be enhanced by the use of high-fidelity medical simulations, possibly due to the task representing actual clinical care involving several patients (Shapiro et al. 2004). However, further research needs to be conducted to see how much simulation training is needed to produce a significant increase in team performance.

As evidenced by certain medical specialties (e.g., emergency medicine), effective inter-professional collaboration and communication is essential to high quality and safe patient care and the prevention of errors. Team training is not limited to physicians, but involves all members of the clinical team, including nurses and other staff. Future directions of research on team training are the development of methods to assess the team process as well as studies to investigate the relationship between team performance and patient safety outcomes.

Conclusions

This paper provides a review of the current issues in evaluation in health professions education and its relationship with student learning and clinical performance, with a focus on medical education. The motivation for effective evaluation serves two major purposes: (1) to provide a scientific foundation for reforming medical curricula that would result in better education and training of competent physicians, hopefully leading to high quality patient care and better patient outcomes; and (2) to provide a learning tool for medical students through the provision of timely and accurate feedback and for assessment of students' progress towards achievement of professional competence. This paper offers several recommendations for curricular change and evaluation, which are empirically and theoretically-based within the framework of the cognitive and learning sciences. We have highlighted the availability of evidence to support curricular reforms to improve student learning and performance, and several areas where further research is needed (e.g., effects of technology and simulation on learning and performance, evaluation of hybrid curricula, research in naturalistic settings). Ultimately, a cognitive perspective on how people think and learn should underlie curricular reform and is necessary to fulfill the mission of medical education.

Acknowledgements The writing of this manuscript was supported in part by funds from the College of Medicine, University of Arizona, Phoenix, in Partnership with Arizona State University.

References

- AAMC (2004). Educating doctors to provide high quality medical care: A vision for medical education in the United States. Report of the Ad Hoc Committee of Deans. Washington, DC.
- AAMC (2005a). Recommendations for clinical skills curricula for undergraduate medical education. Retrieved April 16, 2007, from https://services.aamc.org/Publications/index.cfm?fuseaction=Product.displayForm&prd_id=141&prv_id=165
- AAMC (2005b). Cultural competence education for medical students. Retrieved April 12, 2007, from <http://www.aamc.org/meded/tacct/start.htm>

- AAMC (2006). Implementing the vision: Group on educational affairs responds to the IIME Dean's Committee report. Washington, DC.
- ACGME (1999). ACGME outcome project. General competencies. Retrieved May 3, 2007, from <http://www.acgme.org/outcome/comp/compFull.asp>
- Anderson, J. R. (1983). *The architecture of cognition*. Cambridge, MA: Harvard University Press.
- Anderson, J. R. (1993). *Rules of the mind*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Anderson, J. R. (1996). ACT: A simple theory of complex cognition. *American Psychologist*, *51*, 355–365.
- Anderson, L., & Krathwohl, D. (Eds.) (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. New York: Longman.
- Anderson, J., Reder, L., & Simon, H. (1996). Situated learning and education. *Educational Researcher*, *25*(4), 5–11.
- Anderson, J. R., & Schunn, C. D. (2000). Implications of the ACT-R learning theory: No magic bullets. In R. Glaser (Ed.), *Advances in instructional psychology: Educational design and cognitive science* (Vol. 5, pp. 1–34). Mahwah, NJ: Lawrence Erlbaum Associates.
- Baker, E. L. (2007). Moving to the next generation system design: Integrating cognition, assessment, and learning. CSE Report 706. National Center for Research on Evaluation, Standards, and Student Testing, University of California, Los Angeles.
- Baker, D. P., Gustafson, S., Beaubien, J. M., Salas, E., & Barach, P. (2005). Medical team training programs in health care. In *Advances in patient safety: From research to implementation* (Vol. 4). Rockville, MD: Agency for Healthcare Research and Quality. <http://www.ahrq.gov/qual/advances/>
- Bereiter, C., & Scardamalia, M. (1996). Rethinking learning. In D. R. Olson & N. Torrance (Eds.), *The handbook of education and human development: New models of learning, teaching and schooling* (pp. 485–513). Cambridge, MA: Basil Blackwell.
- Betancourt, J. R. (2003). Cross-cultural medical education: Conceptual approaches and frameworks for evaluation. *Academic Medicine*, *78*, 560–569.
- Blaskiewicz, R. J., Park, R. S., Chibnall, J. T., & Powell, J. K. (2004). The influence of testing context and clinical rotation order on students' OSCE performance. *Academic Medicine*, *79*, 597–601.
- Bloom, B. S. (1956). *The taxonomy of educational objectives*. London: Longman.
- Bond, W. F., Lammers, R. L., Spillane, L. L., Smith-Coggins, R., Fernandez, R., Reznick, M. A., et al. (2007). The use of simulation in emergency medicine: A research agenda. *Academic Emergency Medicine*, *14*, 353–363.
- Bransford, J., Brown, A., & Cocking, R. (2000). *How people learn: Brain, mind, experience, and school* (Expanded ed.). Washington, DC: Commission on Behavioral and Social Sciences and Education/National Academy Press.
- Bransford, J. D., & Schwartz, D. L. (1999). Rethinking transfer: A simple proposal with multiple implications. *Review of Research in Education*, *24*, 61–100.
- Chatterji, M. (2005). Evidence on “what works”: An argument for extended-term mixed-method (ETMM) evaluation designs. *Educational Researcher*, *34*, 14–24.
- Clark, R. C., & Mayer, R. E. (2007). *E-learning and the science of instruction* (2nd ed.). San Francisco: Jossey-Bass.
- Cooper, H., Carlisle, C., Gibbs, T., & Watkins, C. (2001). Developing an evidence base for interdisciplinary learning: A systematic review. *Journal of Advanced Nursing*, *35*, 228–237.
- Detterman, D. K., & Sternberg, R. J. (1993). *Transfer on trial: Intelligence, cognition, and instruction*. Norwood, NJ: Ablex.
- Devitt, J. H., Kurrek, M. M., Cohen, M. M., Fish, K., Fish, P., Noel, A., et al. (1998). Testing internal consistency and construct validity during evaluation of performance in a patient simulator. *Anesthesia & Analgesia*, *86*, 1160–1164.
- Doll, R. C. (1996). *Curriculum improvement* (Rev. ed.). Boston, MA: Allyn and Bacon.
- Dutta, S., Gaba, D., & Krummel, T. M. (2006). To simulate or not to simulate: What is the question? *Annals of Surgery*, *243*, 301–303.
- Epstein, R. M. (2007). Assessment in medical education. *New England Journal of Medicine*, *356*(4), 387–396.
- Ericsson, K. A. (2004). Deliberate practice and the acquisition and maintenance of expert performance in medicine and related domains. *Academic Medicine*, *79*(10 Suppl), S70–S81.
- Ericsson, K. A., Charness, N., Feltovich, P. J., & Hoffman, R. R. (2006). *The Cambridge handbook of expertise and expert performance*. New York: Cambridge University Press.
- Eva, K. W., Cunningham, J. P. W., Reiter, H. I., Keane, D. R., & Norman, G. R. (2004). How can I know what I don't know? Poor self assessment in a well-defined domain. *Advances in Health Sciences Education*, *9*, 211–224.

- Eva, K. W., Neville, A. J., & Norman, G. R. (1998). Exploring the etiology of content specificity: Factors influencing analogic transfer and problem solving. *Academic Medicine*, 73(10 Suppl), S1–S5.
- Feltovich, P. J., Spiro, R. J., & Coulson, R. L. (1993). Learning, teaching, and testing for complex conceptual understanding. In N. Frederiksen, R. J. Mисlevy, & I. I. Bejar (Eds.), *Test theory for a new generation of tests* (pp. 181–217). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Fried, G. M., Feldman, L. S., Vassiliou, M. C., Fraser, S. A., Stanbridge, D., Ghitulescu, G., et al. (2004). Proving the value of simulation in laparoscopic surgery. *Annals of Surgery*, 240, 518–528.
- Friedrich, M. (2002). Risk-free medical training with patient simulators. *Journal of the American Medical Association*, 288, 2808–2812.
- Gaba, D. M. (2004). The future vision of simulation in health care. *Quality and Safety in Health Care*, 13, i2–i10.
- Gallagher, A. G., & Cates, C. U. (2004). Approval of virtual reality training for carotid stenting: What this means for procedural-based medicine. *Journal of the American Medical Association*, 292, 3024–3026.
- Garg, A. X., Adhikari, N. K. J., McDonald, H., Rosas-Arellano, M. P., Devereaux, P. J., Beyene, J., et al. (2005). Effects of computerized clinical decision support systems on practitioner performance and patient outcomes: A systematic review. *Journal of the American Medical Association*, 293, 1223–1238.
- Gelman, R., & Greeno, J. G. (1989). On the nature of competence: Principles of understanding in a domain. In L. B. Resnick (Ed.), *Knowing, learning and instructing: Essays in honor of Robert Glaser* (pp. 125–186). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Gentner, D., Loewenstein, J., & Thompson, L. (2003). Learning and transfer: A general role for analogical encoding. *Journal of Educational Psychology*, 95, 393–408.
- Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive Psychology*, 15, 1–38.
- Gordon, J. A., Brown, D. F. M., & Armstrong, E. G. (2006). Can a simulated critical care encounter accelerate basic science learning among preclinical medical students? A pilot study. *Simulation in Healthcare*, 1, 13–17.
- Gordon, J., Shaffer, D., Raemer, D., Pawlowski, J., Hurford, W., & Cooper, J. (2006). A randomized controlled trial of simulation-based teaching versus traditional instruction in medicine: A pilot study among clinical medical students. *Advances in Health Sciences Education*, 11, 33–39.
- Greeno, J. G. (1998). The situativity of knowing, learning, and research. *American Psychologist*, 53(1), 5–26.
- Greeno, J. G. (2006). Learning in activity. In R. K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (pp. 79–96). New York: Cambridge University Press.
- Gregg, J., & Saha, S. (2006). Losing culture on the way to competence: The use and misuse of culture in medical education. *Academic Medicine*, 81, 542–547.
- Griffith, C. H., Wilson, J. F., Haist, S. A., & Ramsbottom-Lucier, M. (1997). Relationships of how well attending physicians teach to their students' performances and residency choices. *Academic Medicine*, 72, S118–S120.
- Griffith, C. H., Wilson, J. F., Haist, S., & Ramsbottom-Lucier, M. (1998). Do students who work with better housestaff in their medicine clerkships learn more. *Academic Medicine*, 73, S57–S59.
- Guest, C., Regehr, G., & Tiberius, R. (2001). The life long challenge of expertise. *Medical Education*, 35, 78–81.
- Hamdy, H., Prasad, K., Anderson, M. B., Scherpbier, A., Williams, R., Zwierstra, R., et al. (2006). BEME systematic review: Predictive values of measurements obtained in medical schools and future performance in medical practice. *Medical Teacher*, 28, 103–116.
- Hamstra, S. J., Dubrowski, A., & Backstein, D. (2006). Teaching technical skills to surgical residents: A survey of empirical research. *Clinical Orthopaedics and Related Research*, 449, 108–115.
- Hawkes, S., & Santhya, K. G. (2002). Diverse realities: Sexually transmitted infections and HIV in India. *Sexually Transmitted Infections*, 78, 31–39.
- Hobgood, C., Sawning, S., Bowen, J., & Savage, K. (2006). Teaching culturally appropriate care: A review of educational models and methods. *Academic Emergency Medicine*, 13, 1288–1295.
- Holyoak, K. J. (2005). Analogy. In K. J. Holyoak & R. G. Morrison (Eds.), *The Cambridge handbook of thinking and reasoning* (pp. 117–142). New York: Cambridge University Press.
- Howley, L. D. (2004). Performance assessment in medical education: Where we've been and where we're going. *Evaluation in the Health Professions*, 27, 285–303.
- Issenberg, S., McGaghie, W., & Petrusa, E. (2005). Features and uses of high fidelity medical simulations that lead to effective learning: A BEME systematic review. *Medical Teacher*, 27, 10–28.
- Jolly, B. C., Jones, A., Dacre, J. E., Elzubier, M., Kopelman, P., & Hitman, G. (1996). Relationships between students' clinical experiences in introductory clinical courses and their performance on an objective structured clinical examination (OSCE). *Academic Medicine*, 71, 909–916.

- Junger, J., Schafer, S., Roth, C., Schellberg, D., Friedman Ben-David, M., & Nikendei, C. (2005). Effects of basic clinical skills training on objective structured clinical examination performance. *Medical Education, 39*, 1015–1020.
- Kapur, P., & Steadman, R. (1998). Patient simulator competency testing: Ready for takeoff? *Anesthesia and Analgesia, 86*, 1157–1158.
- Kaufman, D. (2003). Applying educational theory in practice. *British Medical Journal, 326*, 213–216.
- Krathwohl, D. R. (2002). A revision of bloom's taxonomy: An overview. *Theory into practice, 41*, 212–218.
- Lave, J. (1988). *Cognition in practice: Mind, mathematics, and culture in everyday life*. Cambridge: Cambridge University Press.
- Lave, J., & Wenger, E. (1991). *Situated cognition: Legitimate peripheral participation*. Cambridge: Cambridge University Press.
- Magee, M. (2006). *State of the field review: Simulation in education. A final report*. Calgary, AB: Alberta Online Learning Consortium.
- Mayer, R. E., Heiser, J., & Lonn, S. (2001). Cognitive constraints on multimedia learning: When presenting more material results in less understanding. *Journal of Educational Psychology, 93*, 187–198.
- Mayer, R. E., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist, 38*(1), 43–52.
- McLaughlin, S. A., Bond, W., Promes, S., & Spillane, L. (2006). The status of human simulation training in emergency medicine residency programs. *Simulation in Healthcare, 1*, 18–21.
- McLaughlin, S. A., Doezema, D., & Sklar, D. P. (2002). Human simulation in emergency medicine training: A model curriculum. *Academic Emergency Medicine, 9*, 1310–1318.
- Metheny, W. P., Espey, E. L., Bienstock, J., Cox, S. M., Erickson, S. S., Goepfert, A. R., et al. (2005). To the point: Medical education reviews evaluation in context: Assessing learners, teachers, and training programs. *American Journal of Obstetrics and Gynecology, 192*, 34–37.
- Morgan, P., & Cleave-Hogg, D. (2002). A worldwide survey of the use of simulation in anesthesia. *Canadian Journal of Anaesthesia, 49*, 659–662.
- Norcini, J. J., Lipner, R. S., & Kimball, H. R. (2002). Certifying examination performance and patient outcomes following acute myocardial infarction. *Medical Education, 36*, 853–859.
- Norcini, J. J., & Mckinley, D. W. (2007). Assessment methods in medical education. *Teaching and Teacher Education, 23*, 239–250.
- Norman, G. R. (2005a). Editorial – Inverting the pyramid. *Advances in Health Sciences Education, 10*, 85–88.
- Norman, G. R. (2005b). Editorial – Beyond PBL. *Advances in Health Sciences Education, 9*, 257–260.
- Norman, G., & Feightner, J. (1981). A comparison of behavior on simulated patients and patient management problems. *Medical Education, 15*, 26–32.
- Patel, V. L., Arocha, J. F., Diermeier, M., How, J., & Mottur-Pilson, C. (2001). Cognitive psychological studies of representation and use of clinical practice guidelines. *International Journal of Medical Informatics, 63*(3), 147–167.
- Patel, V. L., & Dauphinee, W. D. (1984). Return to basic sciences after clinical experience in undergraduate medical training. *Medical Education, 18*(4), 244–248.
- Patel, V. L., Evans, D. A., & Groen, G. J. (1989). Biomedical knowledge and clinical reasoning. In D. A. Evans & V. L. Patel (Eds.), *Cognitive science in medicine: Biomedical modeling* (pp. 53–112). Cambridge, MA: The MIT Press.
- Patel, V. L., Glaser, R., & Arocha, J. F. (2000a). Cognition and expertise: Acquisition of medical competence. *Clinical and Investigative Medicine, 23*(4), 256–260.
- Patel, V. L., Groen, G. J., & Norman, G. R. (1993). Reasoning and instruction in medical curricula. *Cognition and Instruction, 10*(4), 335–378.
- Patel, V. L., Kaufman, D. R., & Arocha, J. F. (2000b). Conceptual change in the biomedical, health sciences domain. In R. Glaser (Ed.), *Advances in instructional psychology: Educational design and cognitive science* (Vol. 5, pp. 329–392). Mahwah, NJ: Lawrence Erlbaum Associates.
- Patel, V. L., Kaufman, D. R., & Magder, S. A. (1991). Causal reasoning about complex physiological concepts in cardiovascular physiology by medical students. *International Journal of Science Education, 13*, 171–185.
- Pellegrini, C. A. (2006). Surgical education in the United States: Navigating the white waters. *Annals of Surgery, 244*, 335–342.
- Posner, G. J., & Rudnitsky, A. D. (1982). *Course design* (2nd ed.). New York: Longman.
- Salomon, G., & Perkins, D. N. (1989). Rocky roads to transfer: Rethinking mechanism of a neglected phenomenon. *Educational Psychologist, 24*, 113–142.
- Schmidt, H. G., Norman, G. R., & Boshuizen, H. P. A. (1990). A cognitive perspective on medical expertise: Theory and implications. *Academic Medicine, 65*, 611–621.

- Seymour, N. E., Gallagher, A. G., Roman, S. A., O'Brien, M. K., Bansal, V. K., Andersen, D. K., et al. (2002). Virtual reality training improves operating room performance: Results of a randomized, double-blinded study. *Annals of Surgery*, *236*, 458–464.
- Shapiro, M. J., Morey, J. C., Small, S. D., Langford, V., Kaylor, C. J., Jagminas, L., et al. (2004). Simulation based teamwork training for emergency department staff: does it improve clinical team performance when added to an existing didactic teamwork curriculum? *Quality and Safety in Health Care*, *13*, 417–421.
- Shumway, J., & Harden, R. (2003). AMEE Guide No. 25: The assessment of learning outcomes for the competent and reflective physician. *Medical Teacher*, *25*, 569–584.
- Snell, L., Tallett, S., Haist, S., Hays, R., Norcini, J., Prince, K., et al. (2000). A review of the evaluation of clinical teaching: New perspectives and challenges. *Medical Education*, *34*, 862–870.
- Spiro, R. J., Feltovich, P. J., Jacobson, M., & Coulson, R. L. (1991). Cognitive flexibility, constructivism, and hypertext: Advanced knowledge acquisition in ill-structured domains. *Educational Technology*, *31*(5), 24–33.
- Steadman, R. H., Coates, W. C., Huang, Y. M., Matevosian, R., Larmon, B. R., McCullough, L., et al. (2006). Simulation-based training is superior to problem-based learning for the acquisition of critical assessment and management skills. *Critical Care Medicine*, *34*, 151–157.
- Suchman, L. A. (1985). *Plans and situated action: The problem of human-machine communication*. Cambridge: Cambridge University Press.
- Sutherland, L. M., Middleton, P. F., Anthony, A., Hamdorf, J., Cregan, P., Scott, D., et al. (2006). Surgical simulation: A systematic review. *Annals of Surgery*, *243*, 291–300.
- Sweller, J., van Merriënboer, J. J., & Paas, F. G. (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, *10*(3), 251–296.
- Tang, T. S., Fantone, J. C., Bozynski, M. E. A., & Adams, B. S. (2002). Implementation and evaluation of an undergraduate sociocultural medicine program. *Academic Medicine*, *77*, 578–585.
- Tervalon, M. (2003). Components of culture in health for medical students' education. *Academic Medicine*, *78*, 570–576.
- van der Vleuten, C. P. M. (1996). The assessment of professional competence: Developments, research and practical implications. *Advances in Health Science Education*, *1*, 41–67.
- Veloski, J., Boex, J. R., Grasberger, M. J., Evans, A., & Wolfson, D. B. (2006). Systematic review of the literature on assessment, feedback and physicians' clinical performance: BEME Guide No. 7. *Medical Teacher*, *28*, 117–128.
- Wallin, C., Meurling, L., Hedman, L., Hedegård, J., & Felländer-Tsai, L. (2007). Target-focused medical emergency team training using a human patient simulator: Effects on behaviour and attitude. *Medical Education*, *41*, 173–180.
- Wang, E. E., & Vozenilek, J. A. (2005). Addressing the systems-based practice core competency: A simulation-based curriculum. *Academic Emergency Medicine*, *12*, 1191–1194.
- Weinberger, S., & Whitcomb, M. (2003). The clinical education of medical students: Report on millennium conferences I & II. Retrieved April 12, 2007, from <http://www.aamc.org/meded/clinicalskills/>