Systematic Review

Evaluating the clinical reasoning of student health professionals in placement and simulation settings: A systematic review

Jennie Brentnall 1,\*, Debbie Thackray 2 and Belinda Judd 1

|  |
| --- |
| **Citation:** Lastname, F.; Lastname, F.; Lastname, F. Title. *Int. J. Environ. Res. Public Health* **2021**, *18*, x. https://doi.org/10.3390/xxxxx  Academic Editor: Firstname Lastname  Received: date  Accepted: date  Published: date  **Publisher’s Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.    **Copyright:** © 2021 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/). |

1 Work Integrated Learning, Faculty of Medicine and Health, The University of Sydney, Australia; [jennie.brentnall@sydney.edu.au](mailto:jennie.brentnall@sydney.edu.au)

2 Physiotherapy, School of Health Sciences, University of Southampton, United Kingdom; [d.thackray@soton.ac.uk](mailto:d.thackray@soton.ac.uk)

1 Work Integrated Learning, Faculty of Medicine and Health, The University of Sydney, Australia; [belinda.judd@sydney.edu.au](mailto:belinda.judd@sydney.edu.au)

**\*** Correspondence: jennie.brentnall@sydney.edu.au

**Abstract:** (1) Background: Clinical reasoning is essential to the effective practice of autonomous health professionals and is therefore an essential capability to develop as students. This review aimed to systematically identify the tools available to health professional educators to evaluate students’ attainment of clinical reasoning capabilities in clinical placement and simulation settings. (2) Methods: A systemic review of seven databases was undertaken. Peer reviewed, English-language publications reporting studies that developed or tested relevant tools were included. Searches included multiple terms related to clinical reasoning and health disciplines. Data regarding each tool’s conceptual basis and constructs evaluated were systematically extracted and analysed. (3) Results: Most of the 61 included papers evaluated students in medical and nursing disciplines, and over half reported on the *Script Concordance Test* or *Lasater Clinical Judgement Rubric*. A range of conceptual frameworks were referenced, though many papers did not reference any framework. (4) Conclusions: Overall, key outcomes highlighted an emphasis on diagnostic reasoning as opposed to management reasoning. Tools were predominantly aligned with individual health disciplines and with limited cross referencing within the field. Future research into clinical reasoning evaluation tools should build on and refer to existing approaches and consider contributions across professional disciplinary divides.

**Keywords:** clinical reasoning; medicine; nursing; allied health; students; assessment and evaluation

1. Introduction

Systemic changes in healthcare are requiring graduates to be better prepared for work in diverse settings, in teams, and addressing increasingly chronic and complex healthcare needs [1]. For this, graduates require competence in clinical reasoning; the process of ‘gathering and synthesising information; generating hypotheses; and formulating a clinical impression, diagnosis, prognosis, treatment, care, and/or management plan’ [2]. It is clinical reasoning that integrates the ‘cognitive, psychomotor and affective skills’ required to be ‘adaptive, iterative, and collaborative’ [3]. Therefore, the more autonomous and responsible the health professional, and the more dynamic and complex the situation (including technological advancements), the greater the need for clinical reasoning [4]. Accordingly, the development and evaluation of clinical reasoning is an increasing focus in health professional education [5-7]. It is also included as an essential graduate attribute in many health professional programs and a competency in many health professional frameworks internationally [7].

Despite the agreed importance of clinical reasoning, there is a lack of consensus on how it is conceptualised and the definitions of related terminology [2-4, 6, 7]. The term clinical reasoning (CR) is often used synonymously with terms such as decision making, critical thinking, problem solving, clinical judgement, and diagnostic reasoning [4, 7]. It is also used as a ‘short-hand’ for a broad concept [8], with variations in what comprises clinical reasoning by health professional discipline and also within disciplines [3]. The lack of consensus, along with the varied use of terminology, limits the advancement of education that will prepare graduates for multi-disciplinary teamwork in dynamic and complex situations inherent to healthcare.

There have been many attempts to define clinical reasoning across the professions and to find methods to teach and assess all the constructs related to the concept. Furthermore, there is a growing interest to know when a health professional student develops clinical reasoning, and there have been several recent systematic reviews of evaluations of the time frame of when students develop their clinical reasoning. Each of these reviews has taken a different focus. Specifically, Carter, Creedy, and Sidebotham [9] reviewed the tools used to measure ‘critical thinking’ development in undergraduate nursing and midwifery education. They identified tools by examining papers that applied an experimental design and measured critical thinking on multiple occasions in undergraduate nursing and midwifery education [9]. Similarly, Macauley and colleagues [10] systematically reviewed evaluations of clinical reasoning that were used as outcome measures, this time in examining simulation programs in any health profession. They also included broad outcome measures that did not necessarily focus on clinical reasoning [e.g., the physical therapy entry level competency assessment, the Assessment of Physiotherapy Practice; 10]. In a recent scoping review, Daniel and colleagues [5] expanded to investigate approaches to clinical reasoning evaluation used in a range of settings (workplace-based, simulation-based, and non-workplace-based) rather than only outcomes-based research, but restricted their review to the evaluation of the clinical reasoning of medical students, residents, and physicians. Between these reviews [5, 9, 10] a wide range of measures have been identified, but only one review has extended beyond medicine and nursing [10]. The focus of these reviews has also generally been on research outcome measures [5, 9]. Yet there are a myriad of approaches to the evaluation of students that are suited to application in different educational contexts, but not necessarily suited to use as research outcome measures [10].

Considering the need of education providers to identify their student health professionals’ proficient development of clinical reasoning in preparation for complex and uncertain work, there remains a need to establish means by which this may be evaluated. There exist dominant theories regarding the development of expertise in clinical reasoning. Namely, *Script Theory* [11, 12] describes the restructuring of knowledge as reasoning is practiced and reinforced with the development of expertise such that novices use knowledge networks to progress through detailed reasoning in a cognitively-demanding process, whereas experts use ‘illness scripts’ to efficiently target information gathering and checking and arrive at a total solution. Alternately, clinical reasoning has been viewed as a skill to which Dreyfus and Dreyfus’ *Model of Skill Development* [13] applies, positing that novices are reliant on rules learned from others and skill progresses through a number of stages of increasing capability to recognise patterns and handle uncertainty through to expertise where solutions are intuitively recognised. These models have informed the conceptualisation of clinical reasoning development in health professions education.

With clinical reasoning being a consequential capability (38) that inherently requires flexible application, there is great interest in its development in experiential simulated and workplace-based settings, and thereby evaluating it as an *applied* skill [14] or capability. Therefore, not a skill suited to evaluation in a classroom setting, and not focussing on purely measuring a student’s cognitive capacity or knowledge base [15]. However, there has not been any research in this area of reasoning across disciplines. Further, no research to date has synthesised the evidence on what the tools of ‘clinical reasoning’ evaluations are purporting to measure, even in medicine and nursing. And as such, may also provide a basis for furthering understanding of the constructs used, and the relationships between them. With these clear gaps, the aim of this review is to systematically identify the tools used to evaluate clinical reasoning and determine the constructs the tools intend to assess.

The following research questions guided this review:

1. What tools have been developed or investigated for evaluating students’ clinical reasoning as applied in clinical education placement and simulation settings of health professional education?
2. What constructs or aspects of clinical reasoning are those tools designed to assess?

2. Materials and Methods

In a systematic approach, potentially eligible studies were identified by searching the databases of CINAHL (via EBSCO), ERIC, EMBASE, Medline and pre-Medline, and PsychInfo (via Ovid), and Proquest Nursing and Allied Health. The search strategy (see Supplemental Table S1) was inclusive of a wide range of allied health disciplines and medicine: audiology, dietetics, exercise physiology, medicine, nursing, occupational therapy, paramedicine, pharmacy, physiotherapy / physical therapy, podiatry, psychology, social work, speech / speech and language pathology / therapy. Students were learners in their primary professional training (i.e., not post-professional continuing development or specialisation). The terms reasoning and the variants critical thinking, judgement, problem solving, and decision making were included, mapped to database subject headings as relevant. Likewise, the search included a range of terms pertaining to evaluation and measurement (e.g., assessment, inventory, test, scale, measure, index, survey, rubric, and so on), adjusted to database subject headings as relevant. Searches were limited to publication dates of 2000- 2018, in English language only, and where possible to peer reviewed sources.

Citations were imported into Covidence (Melbourne, Victoria, Australia) for management. Abstracts, and where necessary full-text papers, were each screened by one author according to the inclusion and exclusion criteria set out in Table 1. To ensure consistency within the team, at each stage papers were screened until several included studies were identified and then the team met to discuss inclusion and exclusion decisions until consensus and consistency were reached. Screening then continued, with uncertainties discussed between the authors in regular meetings.

**Table 1.** Inclusion and exclusion criteria.

|  |  |
| --- | --- |
| **Inclusion criteria** | **Exclusion criteria** |
| Studies meeting ALL the following:   1. Peer reviewed 2. Published 2000-2018 3. Published in English 4. Accessible to these authors 5. Any health profession 6. Pre-registration student education 7. Clinical reasoning or related concepts 8. Primary outcome is to develop or test an evaluation tool | 1. Used but did not develop an evaluation tool 2. Developed or tested an evaluation of a whole range of clinical skills or competencies, even if these included a component on clinical reasoning 3. Evaluation tools only used outside of clinical placement or simulation settings |

This review focused on evaluations used in clinical placement or simulation settings. However, a conservative approach was taken, whereby if the setting in which the student was evaluated was unclear in the paper, but the evaluation was reported as connected to student learning in clinical or simulation settings (i.e., eligible settings), the paper was included.

Data extraction was completed by all authors and included: authorship and year of publication, country or countries in which the study was undertaken, the tool or tools that were studied, the aims, the methods, the disciplines and levels of students and any other participants, the theoretical underpinning of the tool (as stated by the authors), and the construct evaluated. Where papers included multiple evaluation tools, data were extracted for each relevant tool.

To ensure consistency, for each data extraction item, examples were discussed among the team and noted at the top of the data collection table prior to data extraction. Initially data were extracted from two papers by different team members, and these were discussed among the team to reach consensus and clarity. Data from each eligible paper were then extracted by one team member per paper, with all team members working into the common data extraction table where the others’ work was visible. Uncertainties were discussed at regular meetings, with a second team member allocated to read individual papers as required. Finally, all data for this report were reviewed by the first author.

3. Results

The 7,882 records identified in database searches were narrowed to 61 included papers (see Figure 1). Of the 197 papers that appeared to meet the criteria or were unclear from the abstract screening, 135 were excluded at full-text review, predominantly because the evaluation tools were not used in clinical or simulation-based settings (n=46) or because the paper did not describe the development or testing of an evaluation tool (n=41).

7882 records identified in database searches

7825 abstracts screened

196 full-text papers assessed for eligibility

61 papers included

*57 duplicates removed*

*7629 papers excluded based on abstracts*

*135 papers excluded on full-text review, as follows:*

* *98 papers meeting exclusion criteria*

1. *did not develop or test an evaluation tool (n=41)*
2. *tools evaluated a wide range of skills (n=11)*
3. *were not in clinical or SIM settings (n=46)*

* *37 papers otherwise not meeting inclusion criteria*

**Figure 1.** Flow diagram of search.

3.1. Overview of included studies

The majority of the included 61 papers described studies within medicine and nursing, with 28 and 25 papers respectively, plus one paper including participants from both of those professions. The remaining papers were in midwifery (n=3), physical therapy (n=2), occupational therapy (n=1) and pharmacy (n=1). Around half of the papers addressed the development and testing of the *Script Concordance Test* [SCT; n=19; 16] or the *Lasater Clinical Judgement Rubric* [LCJR; n=13; 17] and variants. The SCT is a test for predominantly diagnostic medical scenarios, where the examinee’s answers are scored based on the level of agreement with responses provided by a panel of experts[16]. The LCJR describes performance expectations, as well as language for feedback and assessment of predominantly nursing students’ clinical judgment development in a detailed and developmental rubric [17]. See Supplementary Table S2 for an overview of papers by evaluation tool.

The papers were dominated by studies undertaken in the United States of America (n=24), including five studies of the SCT and seven studies of the LCJR. Other studies were undertaken in Canada (n=8, including five on the SCT), Australia (n=6), France (n=5), Korea (n=5; including three on LCJR variants), and a range of other countries. The distribution of publication years illustrates a trend of increasing publications on this topic over time and particularly since 2015, considering that further papers published in 2018 would not have yet been in the databases at the time of the search.

3.2. Conceptual foundations of the clinical reasoning evaluation tools

The included papers drew from a variety of conceptual frameworks (see Table 2). Representing the topics identified in this review, the included papers are grouped into those stating they evaluate clinical decision making, clinical judgement, clinical reasoning, critical thinking, and situation awareness. The final category are those that do not state a specific construct being measured, though they report measuring clinical reasoning and related constructs in general. Within each group, papers are arranged by the theoretical underpinning that was named in the paper. Evaluation tools for which there were no theoretical underpinnings identified appear at the end of each group.

Even within disciplines, there was a clear lack of agreement regarding critical thinking, with two differing consensus statements in nursing [18, 19] giving rise to several evaluation tools including the *Carter Assessment of Critical Thinking in Midwifery* [20-25] examining students’ critical thinking skills – a construct almost exclusively evaluated for nursing students. Another evaluation of critical thinking skills, through a clinical viva, was used for nursing students [26] and was derived by adapting a nursing competence assessment [27], while another competence assessment [the Physical Therapy Clinical Performance Instrument; 28] was indirectly adapted to create an evaluation of physical therapy interns’ clinical decision making [29]. The one evaluation of critical thinking skills used for medical students [30] was based on a problem solving process [31] and piloted in high fidelity simulation.

Two evaluations utilised context-specific reasoning frameworks, both drawing on mnemonic devices to both guide and evaluate students. An occupational therapy clinical reasoning example [32] applied the *A SECRET* approach [Attention, Sensation, Emotion Regulation, Culture, Relationships, Environment, and Task; 33]. In an example evaluating medical students’ clinical documentation [34], the IDEA framework (Interpretive summary, Differential diagnosis, Explanation of reasoning, and Alternative diagnosis with explanation) was combined with RIME descriptions [Reporter, Interpreter, Manager, Educator; 35].

Contrasting in specificity with the context-specific frameworks was the consideration of broad cognitive frameworks and abilities. Script theory [11, 12], which posits that expert clinical reasoning largely draws upon patterns, was used as the foundation of the *Script Concordance Test* [16] and variants including a multiple choice examination [36] and written ‘think aloud’ test [37]. Situation awareness [38] was evaluated for nursing students [39]. Drawing upon an even more general foundation, a critical thinking skills evaluation [40] used for nursing students was based on Benner’s levels of nursing expertise, where students are thought to move through five levels of increasing proficiency; novice, advanced beginner, competent, proficient, and expert [41] and *Bloom’s Taxonomy of Educational Objectives*, where hierarchical models are used to classify educational learning objectives by levels of complexity and specificity[42]. Similarly, a clinical reasoning evaluation used for physical therapy students [43] was based on multiple sources of knowledge regarding clinical reasoning with evaluations based on the *Revised Bloom’s Taxonomy of Educational Objectives* [44] and the *Dreyfus Model of Skill Acquisition* [13] which describes skill development through instruction and experience of five developmental stages from novice to mastery. In another instance, authors reported the complementary use of a clinical reasoning model [45] and social cognitive theory [46, 47], which considers that an individual's thoughts and feelings as well as the social environment affect their own behaviour, to derive an evaluation of anxiety and confidence in clinical decision making [48].

Finally, across all the included papers, published models of clinical judgement [49], clinical decision making, and clinical reasoning processes [45, 50, 51] were cited as directly or indirectly underpinning the *Lasater Clinical Judgement Rubric* [17] and variants [52-56], as well as several evaluations of clinical decision making [48] and clinical reasoning [57, 58]. These represented the most direct link between conceptual models and student evaluations.

Collectively, the identified frameworks represent a broad spectrum of the main constructs deemed important and necessary for the development of clinical reasoning. However, inconsistencies in agreement about the underpinning conceptual framework for clinical reasoning was evident. Further, for 11 evaluation tools adopted for medical and nursing students no framework was specified [59-69]. Many of these evaluation tools were described as rubrics, examinations, and objective structured clinical examinations (OSCEs), and the authors did not clearly ascribe the construct being assessed to be of critical thinking, clinical judgement, clinical decision making, or clinical reasoning. It would appear that if a conceptual framework for clinical reasoning is used, then the specific constructs are identified and built into the assessment tool whereas those without a conceptual framework, it was found that the constructs of the evaluation tools are often only identified in general terms. These tools may, for example, purport to represent ‘clinical reasoning’ generically by way of a proxy activity such as clinical documentation.

**Table 2.** Constructs measured in evaluations of students’ clinical reasoning.

| **Construct** | **Theoretical Underpinning** | **Evaluation Tool or Measure (Discipline of Student)** |
| --- | --- | --- |
| **Clinical Decision Making** | None stated (an adaptation of the Physical Therapist Clinical Performance Instrument) [28] | Clinical Decision Making Survey Tool (Physical Therapy) [29] |
| Not stated | Surgical Decision Making Rating Scale (Medicine) [59] |
| **Clinical Judgement** | Clinical Judgement Model [49] | Lasater Clinical Judgement Rubric (Nursing) [17, 70-75] |
| Lasater Clinical Judgement Rubric – Korean version (Nursing) [52] |
| Lasater Clinical Judgement Rubric – Dutch version (Nursing) [53] |
| Virtual Patient Lasater Clinical Judgement Rubric (Nursing) [54] |
| Scenario-specific Assessment Tool for Febrile Infant Care Simulation (adaptation of the Lasater Clinical Judgement Rubric; Nursing) [55] |
| Simulation Evaluation Tool (an adaptation of the Lasater Clinical Judgement Rubric) (Nursing) [56] |
| **Clinical Reasoning** | “A SECRET” reasoning approach [33] | A SECRET Assessment (Occupational Therapy) [32] |
| Clinical Reasoning Process model [51] | Clinical Reasoning Evaluation Simulation Tool (CREST) (Nursing) [57] |
| Nurses Clinical Reasoning Scale (Nursing) [58] |
| IDEA Framework, structural semantics and RIME [35] | IDEAs Assessment Tool (Medicine) [34] |
| Novice Clinical Reasoning Model [45] and Social cognitive theory [46, 47] | Nursing Anxiety and Self-Confidence with Clinical Decision Making (NASC-CDM; Nursing) [48] |
| Outcome Present State Test Model [50] | Outcome Present State Test (OPT; Nursing) [76] |
| Revised Bloom’s Taxonomy [44] and Dreyfus Model [13] | Clinical Reasoning Grading Rubric (Physical Therapy) [43] |
| Script Theories [11, 12] | Multiple Choice Question Exam (Medicine) [36]\* |
| Script Concordance Test   (Medicine) [36]\*, [77-88], [89]\*, [90]\*  (Nursing) [91], [90]\*  (Pharmacy) [92] |
| Script Concordance Test with Think Aloud (Medicine) [37], [89]\* |
| Not stated | Clinical Reasoning Problems Test (Medicine) [60], [89]\* |
| **Critical Thinking** | Benner’s [41] levels of nursing experience, and Bloom’s [42] cognitive domains | Clark Simulation Evaluation Rubric (Nursing) [40] |
| Consensus dimensions of critical thinking in nursing [18] | Carter Assessment of Critical Thinking in Midwifery (Preceptor / Mentor Version) (Midwifery) [21], [23]\* |
| Carter Assessment of Critical Thinking in Midwifery (Student Self-Rating Version) (Midwifery) [22], [23]\* |
| Carter Assessment of Critical Thinking in Midwifery (Reflective Writing) (Midwifery) [23]\* |
| Rubric for assessing critical thinking dimensions (Nursing) [20] |
| Expert Consensus on Critical Thinking [19] | Critical Thinking Self-Reflection Tool (Nursing) [24] |
| Yoon’s Critical Thinking Tool (Nursing) [25] |
| IDEAS five-step critical thinking problem solving process [31] | Critical Thinking Skills Rating Instrument (CTSRI; Medicine) [30] |
| Structured Observation of and Assessment of Practice [27] | Clinical Viva (Nursing) [26] |
| **Situation Awareness** | Situation Awareness [38] | Situation Awareness Global Assessment Technique (SAGAT; Nursing) [39] |
| **Not Specified** | Not stated | Clinical Performance Examination (CPX; Medicine) [61] |
| Computer-based Case Simulation (CCS; DxR Clinician Software; Medicine) [63] |
| Exam formats: Context-rich single best answer versus key feature problems (Medicine) [62] |
| Exam formats: Extended matching questions, with think aloud (Medicine) [64] |
| Interactive Simulation of Patients Objective Structured Clinical Examination (OSCE) Station (Medicine) [65] |
| Objective Structured Clinical Examination (OSCE) Note Writing Station (Medicine) [66] |
| Reflective Thinking Instrument (Nursing) [67] |
| Virtual Patient Case Patient Summary Statement Rubric (Medicine) [68] |
| Virtual Patient Case Procedural Rubric and Semantic Rubric (Medicine) [69] |

4. Discussion

This review has identified numerous tools used to evaluate clinical reasoning and related constructs in placements and simulation in health professional education. Of these tools, the *Script Concordance Test* [16] and *Lasater Clinical Judgement Rubric* [17] were prominent in the literature. The diversity of additional tools identified from searches using a range of terms provides educators with a variety of options for student evaluation in these situations. These tools encompass a spread of approaches along the ‘continuum of authenticity’ [5], given our inclusion of evaluation tools described as being conducted during, or associated with, clinical placement or simulation settings, even if not explicitly of workplace performance. However, there is a lack of cross-referencing between tools and constructs identified in this review, and evidence of continued development almost exclusively within discipline boundaries. From identifying these tools and their conceptual foundations, we present *four* key implications for further discussion.

4.1. There remains incosistent use of terminology around clinical reasoning

Unsurprisingly given previous reports [2-4, 6, 7], there is a lack of consistency in the application of terminology to name the constructs being assessed. To advance both practice and research, there remains a need to clarify both terminology and constructs, and to do so in ways that will enable better understanding of how each profession contributes similarly and differently to clinical reasoning in the practice of both diagnosis and management [3, 7, 93]. Communicating clearly about, and reconciling, conceptual and theoretical frameworks will be required to advance this cause [6], in the literature more broadly and in reference to the evaluation of clinical reasoning and related constructs specifically.

4.2. The tools each have limited application and minimal multi-disciplinary crossover

This review has also highlighted that the evidence to support the use of evaluation tools, or make choices between them, is in most cases limited for these learning contexts. The *Script Concordance Test* [16] and *Lasater Clinical Judgement Rubric* [17] clearly dominated the published work, but there were many more tools reported on with very limited interconnections made between them or even cross-referencing of research. It is striking that there was no overlap with measures of critical thinking identified in a previous systematic review of studies with experimental designs [9], and limited overlap with those identified in a systematic review of a broad range of measures used to assess simulation outcomes [10]. In part, the development of some of the tools in this review was in response to the need for appropriate outcome measures that prior reviews in specific circumstances have identified, but overwhelmingly the evaluation tools in this review have been developed and tested in isolation of each other and with limited subsequent application. The few papers per available tool offer limited evidence for evaluation of health professional students in clinical and simulated placements.

Most tools were also only used for one professional discipline of student, usually nursing or medicine, as has been the focus of prior reviews [5, 9]. The *Script Concordance Test* has been used predominantly for medical specialties, though also in one study for pharmacy students and two for nursing students. However, it should be noted that the case vignettes for the *Script Concordance Test* were developed and calibrated against content expert decisions for each instance completely independently. Even single studies, where participants were drawn from multiple disciplines or medical specialties [88, 90] used multiple instances of the test rather than direct comparison. As a consequence, there remains no identified avenues of considering how students from differing disciplines engage in clinical reasoning regarding the same patient scenario, with implications for teaching and evaluation of constructs [7], and also for understanding how healthcare teams may work together. If disciplines don’t share a common understanding of the conceptual framework or constructs of clinical reasoning, this may impact on and limit interprofessional learning and collaborative clinical judgements in multidisciplinary teams.

Similar discipline differentiations are visible when considering the description of the constructs that the various identified tools purport to measure. So even within the same discipline there is substantial variance. ‘Critical thinking skills’ have largely been evaluated in clinical placements or simulation in nursing and midwifery. In many cases these have been developed from consensus statements from different regions [18, 19], with varied evaluation tools even from the same consensus statements. ‘Clinical judgement’ was also named as the construct examined in studies in nursing, following the *Clinical Judgement Model* [49] and *Lasater Clinical Judgement Rubric* [17]. ‘Clinical decision making’ was named as the construct evaluated in one study each in medicine, nursing and physical therapy, with limited theory underpinning this choice. Collectively these three constructs appear to represent the more specific of those named in this review.

The tools reported to evaluate ‘clinical reasoning’ were more widespread across disciplines and theoretical foundations, which is consistent with the diverse use of this terminology [7]. Included among these were evaluations of the application of specific frameworks for ‘reasoning’ that were not themselves models of clinical reasoning [32, 34], diagnostic processes as in the *Script Concordance Test* and others [34, 76], and therapeutic processes [57]. Finally, there was a group of papers that examined diverse evaluation tools of relevance to the topic, but without clearly setting out the origin of the construct of interest, sometimes not even clearly naming a construct at all (ref). Again, this is reflective of the usage of terminology, and reinforces the need to make explicit intended meanings [7]. If health professionals can speak the same language when considering CR pedagogically and clinically, the impacts may be seen on student learning, collaborative decision making and patient care.

4.3. There are limited findings applicable to allied health

Somewhat surprisingly, there were only four papers (of the total of 61) that considered tools for the evaluation of clinical reasoning in disciplines beyond nursing and medicine (allied health). The importance of clinical reasoning in these disciplines is similarly critical, with many allied health disciplines being primary care practitioners. Clinical reasoning is also important for patient management decisions in all these professions, with the majority also working in multidisciplinary teams. With the lack of research, it is unclear if findings on existing tools are applicable to allied health student development. Alongside the inconsistent use of terminology, this may impede multi-disciplinary learning and collaborative team-based patient care.

4.4. Evaluation tools reported in the literature represent two contrasting objectives

The two differing objectives of clinical reasoning measurement represented by the tools in this review, represent contrasts in emphasis on diagnostic versus management considerations in reasoning [93]. When comparing the tools side-by-side, it is apparent that evaluating students’ clinical reasoning is complex. The tools identified do not, individually or even between them, represent the full complexity of the construct of reasoning set out in the literature [4, 5, 93]. Some approaches to student evaluation emphasise cognitive functioning, and perhaps specificity and objectivity, as do standardised psychometric measures of critical thinking abilities that have been used in non-clinical settings of health professional education [10]. Other approaches to student evaluation favour comprehensive coverage in practice situations. Each takes different approaches to considering how students manage inherently dynamic healthcare situations. Given none is complete, educators must be mindful of how the application of different constructs when evaluating students’ learning can lead to different interpretations.

4.4.1. Tools to assess the development of diagnostic reasoning

A key objective of assessment in a subset of the papers included in this review is to examine the congruence between students’ reasoning outcomes and those of deemed experts. This is evident predominantly in instances of the *Script Concordance Test*. This objective necessitates the identification of a point of clear comparison, which can restrict the aspects and applications of clinical reasoning that can be evaluated. Whereas definitions of clinical reasoning incorporate the whole therapeutic process [2, 4], the focus from this objective is typically on diagnosis and treatment decisions where agreement can be objectified. The expectation is that students start out using limited network approaches in knowledge organisation and build refined knowledge scripts with experience [94]. This can be a useful way to consider the development of expertise and evaluations using this approach, while usually paper-based and closed-response, are able to track shifts in respondents’ thinking as new data informing clinical reasoning is provided. However, the domains covered need to be married with complementary approaches or comprehensive evaluation to encapsulate all components of clinical reasoning [5].

4.4.2. Tools to judge the quality of performance as a reflection of reasoning processes

The alternate objective of assessment in the papers in this review is to judge the quality of performance of steps that apply throughout a process of therapeutic patient management. Such reasoning processes are clearly set out in models such as the *Clinical Judgement Model* [49] and *Clinical Reasoning Process* [51]. This gives rise to the labelling of this construct as *clinical judgement* in tools such as the *Lasater Clinical Judgement Rubric* [17] while others [57, 58] term something very similar clinical reasoning. These instruments address multiple clinical reasoning components beyond diagnosis or scenario planning, which is more consistent with broad definitions of clinical reasoning [2, 4, 5]. They typically make use of the authentic and dynamic situations in simulation and clinical placements, relying on extended and preferably multiple observations of student reasoning performance with assessor judgement [5]. However, this approach may introduce the risk of misjudgement by assessors if using observable performance outputs (behaviours) to infer cognitive and affective elements of reasoning. It may also be difficult to distinguish between the cognitive, psychomotor and affective skills in reasoning, and broader therapeutic or technical skills. For example, the *Lasater Clinical Judgement Rubric* includes ‘calm, confident manner’, ‘clear communication’, and ‘being skilful’ among the elements of clinical judgement [17]. Evaluations using this approach might therefore be more holistic than specific and sensitive to clinical reasoning and thereby student development.

5. Limitations

This systematic review sought to identify the range of tools reported in the literature for educators to consider in evaluating the clinical reasoning or related abilities of allied health and medical students in simulated or placement settings. This represents only the formally published, English-language literature on the topic, subject to publication bias and limited international representation, particularly as grey literature searching and secondary search strategies were not used (e.g., reference list and citation tracking, or contacting authors). Author reports were used rather than independent analysis of evaluation tools and methods, given the full content of many tools were not available, and thus variation in the use of constructs and terms was visible but not resolved in this study. None-the-less, this review has been able to present a broad overview with respect to the inclusions of disciplines of the students, while focusing on the evaluation of clinical reasoning and related constructs specifically.

6. Conclusions

This study has identified a significant number of tools used to evaluate clinical reasoning and related constructs in placement and simulation settings in health professional education. There is a lack of cross-referencing between tools and constructs identified in this review, and evidence of continued development only within discipline boundaries. Future research into clinical reasoning evaluation tools should build on and reference existing approaches and consider contributions across professional disciplinary divides. Research is needed to develop, test, and incorporate student evaluations that are applicable to outcome measurement in research studies in order to understand students’ performance of this essential capability and how to support its development. A larger evidence base than was identified for most tools in this review is required for that purpose, with attention to research quality. Repeated measures and longitudinal perspectives capturing students’ reasoning development are specifically required, as are workplace-based approaches [14]. By connecting and expanding this body of work, it will be possible to more clearly identify contributors to students’ learning and their attainment of threshold skills. Clearly more work is required to sequence the development of clinical reasoning by standardising the use of terminology and constructs, and considering tool design that can monitor the developmental progression of clinical reasoning progression with applicability across health professions.

**Supplementary Materials:** The following are available online at www.mdpi.com/xxx/s1, Table S1: search strategy, Table S2: tools evaluating students’ clinical reasoning.

**Author Contributions:** All authors contributed to the conceptualization, methodological design, screening and data extraction as described in the paper, data analysis and conduct of the research, and the drafting and review of the publication. All authors have read and agreed to the published version of the manuscript. Preliminary findings of this review have previously been presented at the *8th* *International Clinical Skills Conference*, May 2019, Prato, Italy*.*

**Funding:** The first and third authors have been supported in part by the Special Studies Program of their institution. The second author has been supported in part by a Global Research Initiator Award from their institution.

**Data Availability Statement:** The included papers are referenced within, and data extraction tables are included within and in supplementary materials. A full list of identified references including all those excluded from the study is available from the corresponding author on request.

**Acknowledgments:** In this section, you can acknowledge any support given which is not covered by the author contribution or funding sections. This may include administrative and technical support, or donations in kind (e.g., materials used for experiments).

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

**Supplementary materials (included here for review)**

**Supplementary Table 1.** Search strategy.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Health profession | AND  Student | AND  Reasoning | AND  Tool Development and Testing |
| CINAHL (via EBSCO) | “allied health profession\*” or “allied health therap\*” or “physiotherap\*” or “physical therap\*” or “occupational therap\*” or “social work\*” or “speech patholog\*” or “speech therap\*” or “speech language\*” or “dietician\*” or “dietetic\*” or “podiatr\*” or “audiolog\*” or “psychol\*” or “pharma\*” or “exercise physiolog\*” or “paramedic\*” or “nurs\*” or “medical student\*” | “student” | “reasoning” or “critical thinking” or “judgement” or “problem solving” or “decision making” | “inventory” or “assessment” or “evaluation” or “rubric” or “test” or “scale” or “measure” or “index” or “examination” or “survey” or “questionnaire” or  (MH "Educational Measurement+") |
| EMBASE, ERIC, Medline,  Premedline and PsychINFO (via Ovid) | allied health profession\*.mp or allied health therap\*.mp or physiotherap\*.mp or physical therap\*.mp or occupational therap\*.mp or social work\*.mp or speech patholog\*.mp or speech therap\*.mp or speech language patholog\*.mp or speech language therap\*.mp or dietician.mp or dietetic\*.mp or podiatr\*.mp or audiolog\*.mp or psychol\*.mp or pharma\*.mp or exercise physiolog\*.mp or paramedic\*.mp or nurs\*.mp or medical student\*.mp | Student.mp | reasoning.mp or critical thinking.mp or judgement.mp or problem solving/ or decision making/ | Inventory.mp or Assessment.mp or Evaluation.mp or Rubric.mp or Test.mp or Scale.mp or Measure.mp or Index.mp or Examination.mp or "Surveys and questionnaires"/ or Educational measurement/ |
| Proquest Nursing and Allied Health | noft(“allied health profession\*” or “allied health therap\*” or “physiotherap\*” or “physical therap\*” or “occupational therap\*” or “social work\*” or “speech patholog\*” or “speech therap\*” or “speech language\*” or “dietician\*” or “dietetic\*” or “podiatr\*” or “audiolog\*” or “psychol\*” or “pharma\*” or “exercise physiolog\*” or “paramedic\*” or “nurs\*” or “medical student\*”) | noft(student) | noft(“reasoning” or “critical thinking” or “judgement” or “problem solving” or “decision making”) | noft(“inventory” or “assessment” or “evaluation” or “rubric” or “test” or “scale” or “measure” or “index” or “examination” or “survey” or “questionnaire”) |

**Supplementary Table 2.** Tools evaluating students’ clinical reasoning.

| **Evaluation Tool or Measure** | **Paper First Author1, Year** | **Country of Study** | **Discipline of Learner** |
| --- | --- | --- | --- |
| A SECRET Assessment | Gee, 2017 [32] | USA | Occupational Therapy |
| Clark Simulation Evaluation Rubric | Gantt, 2010 [40] | USA | Nursing |
| Clinical Reasoning Evaluation Simulation Tool (CREST) | Liaw, 2018 [57] | Australia (+ international) | Nursing |
| Carter Assessment of Critical Thinking in Midwifery (Preceptor / Mentor Version) | Carter, 2016 [21]  Carter, 2018[23]\* | Australia | Midwifery |
| Carter Assessment of Critical Thinking in Midwifery (Student Self-Rating Version) | Carter, 2017 [22]  Carter, 2018 [23]\* | Australia | Midwifery |
| Carter Assessment of Critical Thinking in Midwifery (Reflective Writing) | Carter, 2018 [23]\* | Australia | Midwifery |
| Clinical Decision Making Survey Tool | Brudvig, 2017 [29] | USA | Physical therapy |
| Clinical Performance Examination (CPX) | Im, 2016 [61] | Korea | Medicine (n.s.) |
| Clinical Reasoning Problems Test | Derakhshandeh, 2018 [60] | Iran | Medicine (Cardiology) |
|  | Groves, 2013 [89]\* | Australia | Medicine (GP) |
| Computer-based Case Simulation (CCS; DxR Clinician Software) | Fida, 2015 [63] | Bahrain | Medicine (n.s.) |
| Critical Thinking Self-Reflection Tool | Cise, 2004 [24] | USA | Nursing |
| Critical Thinking Skills Rating Instrument (CTSRI) | Nguyen, 2017 [30] | USA | Medicine (n.s.) |
| Exam format: Context-rich single best answer versus key feature problems | Huwendiek, 2017 [62] | Germany | Medicine (Paediatrics) |
| Exam format: Extended matching questions (with think aloud) | Beullens, 2005 [64] | Belgium | Medicine (Various) |
| Exam format: Multiple Choice Question | Kelly, 2012 [36]\* | USA | Medicine (Internal) |
| IDEAs Assessment Tool | Baker, 2015 [34] | USA | Medicine (n.s.) |
| Lasater Clinical Judgement Rubric | Adamson, 2012 [71]  Adamson, 2016 [71]  Ashcraft, 2013 [72]  Bussard, 2018 [73]  Lasater, 2007 [17]  Manetti, 2018 [74]  Strickland, 2017 [75] | USA | Nursing |
| Román-Cereto, 2018 [95] | Spain | Nursing |
| Lasater Clinical Judgement Rubric – Dutch version | Vreugdenhil, 2018 [53] | The Netherlands | Nursing |
| Lasater Clinical Judgement Rubric – Korean version | Shin, Park & Shim, 2015 [52] | Korea | Nursing |
| Lasater Clinical Judgement Rubric – Virtual Patient | Georg, 2018 [54] | Sweden | Nursing |
| Lasater Clinical Judgement Rubric adaptation: Scenario-specific Assessment Tool for Febrile Infant Care Simulation | Shin, Shim, Lee & Quinn, 2014 [55] | Korea | Nursing |
| Lasater Clinical Judgement Rubric adaptation: Simulation Evaluation Tool | Kim, 2016 [56] | Korea | Nursing |
| Nurses Clinical Reasoning Scale | Liou, 2016 [58] | Taiwan | Nursing |
| Nursing Anxiety and Self-Confidence with Clinical Decision Making (NASC-CDM) | White, 2014 [48] | USA | Nursing |
| Objective Structured Clinical Examination (OSCE) Station with Interactive Simulation of Patients | Courteille, 2008 [65] | Sweden | Medicine (Surgery) |
| Objective Structured Clinical Examination (OSCE) Station – Note Writing | Berger, 2012 [66] | USA | Medicine (n.s.) |
| Outcome Present State Test (OPT) | Kautz, 2009 [76] | USA | Nursing |
| Reflective Thinking Instrument | Tutticci, 2016 [67] | Australia | Nursing |
| Rubric, Clinical Reasoning Grading | Furze, 2015 [43] | USA | Physical therapy |
| Rubric, Critical Thinking Dimensions | Allen, 2004 [20] | USA | Nursing |
| Rubric, Virtual Patient Case Patient Summary Statement | Smith, 2016 [68] | USA | Medicine (Various) |
| Rubric, Virtual Patient Case Procedural Rubric and Semantic | Fleiszer, 2018 [69] | Canada | Medicine (Surgery) |
| Surgical Decision Making Rating Scale | Chatterjee, 2009 [59] | Canada | Medicine (Urology) |
| Script Concordance Test | Amini, 2017 [77] | Iran | Medicine (Internal) |
| Boulouffe, 2014 [78] | Canada | Medicine (Emergency) |
| Dawson, 2014 [91] | USA | Nursing |
| Funk, 2017 [92] | USA | Pharmacy |
| Gagnon, 2009 [90] | Canada | Medicine (Radiation Oncology & Paediatrics) & Nursing |
| Goos, 2016 [79] | Germany | Medicine (Surgery) |
| Groves, 2013 [89]\* | Australia | Medicine (GP) |
| Humbert, 2011 [80] | USA | Medicine (Emergency) |
| Kania, 2011 [81] | France | Medicine (ENT) |
| Kazour, 2017 [82] | France | Medicine (Psychiatry) |
| Kelly, 2012 [36]\* | USA | Medicine (Internal) |
| Lambert, 2009 [83] | Canada | Medicine (Radiation Oncology) |
| Lubarsky, 2009 [96] | Canada | Medicine (Neurology) |
| Ruiz, 2010 [84] | USA | Medicine (Internal/ Geriatrics) |
| Sibert, Charlin et al., 2002 [85]  Sibert, Darmoni et al., 2006 [86] | France | Medicine (Urology) |
| Subra, 2017 [87] | France | Medicine (GP) |
| Wan, 2018 [88] | Australia | Medicine (Various) |
| Script Concordance Test with Think Aloud | Power, 2017 [37] | Canada | Medicine (Paediatrics) |
| Groves, 2013 [89]\* | Australia | Medicine (GP) |
| Situation Awareness Global Assessment Technique (SAGAT) | Lavoie, 2016 [39] | Canada | Nursing |
| Viva, clinical | Roberts, 2013 [26] | UK | Nursing |
| Yoon’s Critical Thinking Tool | Shin, Park & Kim, 2015 [25] | Korea | Nursing |

References

1. McLaughlin, J.E., et al., *A qualitative review of the design thinking framework in health professions education.* BMC Medical Education, 2019. **19**: p. 98.

2. Young, M., et al., *Drawing boundaries: The difficulty in defining clinical reasoning.* Academic Medicine, 2018. **93**(7): p. 990-995.

3. Huhn, K., et al., *Clinical reasoning in physical therapy: A concept analysis.* Physical Therapy, 2019. **99**(4): p. 440-456.

4. Simmons, B., *Clinical reasoning: Concept analysis.* Journal of Advanced Nursing, 2010. **66**(5): p. 1151-8.

5. Daniel, M., et al., *Clinical reasoning assessment methods: A scoping review and practical guidance.* Academic Medicine, 2019. **94**(6): p. 902-912.

6. Durning, S.J., et al., *Clarifying assumptions to enhance our understanding and assessment of clinical reasoning.* Academic Medicine, 2013. **88**(4): p. 442-448.

7. Young, M., et al., *The terminology of clinical reasoning in health professions education: Implications and considerations.* Medical Teacher, 2019. **41**(11): p. 1277-1284.

8. Young, M., *Crystallizations of constructs: Lessons learned from a literature review.* Perspectives in Medical Education, 2018. **Supplement 1**: p. 21-23.

9. Carter, A.G., D.K. Creedy, and M. Sidebotham, *Evaluation of tools used to measure critical thinking development in nursing and midwifery undergraduate students: A systematic review.* Nurse Education Today, 2015. **35**(7): p. 864-874.

10. Macauley, K., et al., *Systematic review of assessments that evaluate clinical decision making, clinical reasoning, and critical thinking changes after simulation participation.* Journal of Physical Therapy Education, 2017. **31**(4): p. 64-75.

11. Charlin, B., J. Tardif, and H.P.A. Boshuizen, *Scripts and medical diagnostic knowledge: Theory and applications for clinical reasoning instruction and research.* Academic Medicine, 2000. **75**(2): p. 182-190.

12. Schmidt, H.G., G.R. Norman, and H.P.A. Boshuizen, *A cognitive perspective on medical expertise: Theory and implications.* Academic Medicine, 1990. **65**(10): p. 611-321.

13. Dreyfus, S.E. and H.L. Dreyfus, *A five-stage model of the mental activities involved in directed skill acquisition*. 1980, Operations Research Center, University of California: Berkley, CA.

14. Kononowicz, A.A., et al., *The need for longitudinal clinical reasoning teaching and assessment: Results of an international survey.* Medical Teacher, 2020: p. 1-6.

15. Renic, J., et al., *Clinical reasoning performance assessment: Using situated cognition theory as a conceptual framework.* Diagnosis, 2020. **7**(3): p. 241-249.

16. Charlin, B., et al., *The Script Concordance Test: A tool to assess the reflective clinician.* Teaching and Learning in Medicine, 2000. **12**(4): p. 189-95.

17. Lasater, K., *Clinical judgment development: Using simulation to create an assessment rubric.* Journal of Nursing Education, 2007. **46**(11): p. 496-503.

18. Scheffer, B.K. and M.G. Rubenfeld, *A consensus statement on critical thinking in nursing.* Journal of Nursing Education, 2000. **39**(8): p. 352-359.

19. American Philosophical Association, *Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction*. 1990, Millbrae, CA: California Academic Press.

20. Allen, G.D., M.G. Rubenfeld, and B.K. Scheffer, *Reliability of assessment of critical thinking.* Journal of Professional Nursing, 2004. **20**(1): p. 15-22.

21. Carter, A.G., D.K. Creedy, and M. Sidebotham, *Development and psychometric testing of the Carter Assessment of Critical Thinking in Midwifery (Preceptor/Mentor version).* Midwifery, 2016. **34**: p. 141-149.

22. Carter, A.G., D.K. Creedy, and M. Sidebotham, *Critical thinking skills in midwifery practice: Development of a self-assessment tool for students.* Midwifery, 2017. **50**: p. 184-192.

23. Carter, A.G., D.K. Creedy, and M. Sidebotham, *Measuring critical thinking in pre-registration midwifery students: A multi-method approach.* Nurse Education Today, 2018. **61**: p. 169-174.

24. Cise, J.S., C.S. Wilson, and M.J. Thie, *A qualitative tool for critical thinking skill development.* Nurse Educator, 2004. **29**(4): p. 147-151.

25. Shin, H., C.G. Park, and H. Kim, *Validation of Yoon's Critical Thinking Disposition Instrument.* Asian Nursing Research, 2015. **9**(4): p. 342-348.

26. Roberts, D., *The clinical viva: An assessment of clinical thinking.* Nurse Education Today, 2013. **33**(4): p. 402-406.

27. Levett-Jones, T., et al., *Implementing a clinical competency assessment model that promotes critical reflection and ensures nursing graduates' readiness for practice.* Nurse Education in Practice, 2011. **11**(1): p. 64-69.

28. Roach, K.E., et al., *Validation of the revised physical therapist clinical performance instrument (PT CPI): Version 2006.* Physical Therapy, 2012. **92**(1): p. 416-428.

29. Brudvig, T.J., K. Macauley, and N. Segal, *Measuring clinical decision-making and clinical skills in DPT students across a curriculum: Validating a new survey tool.* Journal of Allied Health, 2017. **46**(1): p. 21-26.

30. Nguyen, K., et al., *Developing a tool for observing group critical thinking skills in first-year medical students: a pilot study using physiology-based, high-fidelity patient simulations.* Advances in Physiology Education, 2017. **41**(4): p. 604-611.

31. Facione, P.A., *Critical thinking: What it is and why it counts*. 2015, Hermosa Beach, CA: Measured Reasons.

32. Gee, B.M., et al., *The development of a measurement tool evaluating knowledge related to sensory processing among graduate occupational therapy students: A process description.* Occupational Therapy International, 2017. **2017**.

33. Bialer, D.S. and L.J. Miller, *No longer A SECRET: Unique common sense strategies for children with sensory or motor challenges*. 2011, Arlington, TX: Sensory World.

34. Baker, E.A., et al., *The IDEA assessment tool: Assessing the reporting, diagnostic reasoning, and decision-making skills demonstrated in medical students' hospital admission notes.* Teaching and Learning in Medicine, 2015. **27**(2): p. 163-173.

35. Pangaro, L., *A new vocabulary and other innovations for improving descriptive in-training evaluations.* Academic Medicine, 1999. **74**: p. 1203-1207.

36. Kelly, W., S. Durning, and G. Denton, *Comparing a script concordance examination to a multiple-choice examination on a core internal medicine clerkship.* Teaching and Learning in Medicine, 2012. **24**(3): p. 187-193.

37. Power, A., J.F. Lemay, and S. Cooke, *Justify your answer: The role of written think aloud in script concordance testing.* Teaching and Learning in Medicine, 2017. **29**(1): p. 59-67.

38. Endsley, M.R., *Theoretical underpinning of situation awareness: A critical review*, in *Situation awareness analysis and measurement*, M.R. Endsley and D.J. Garland, Editors. 2000, Lawrence Erlbaum: Mahwah, NJ. p. 3-32.

39. Lavoie, P., S. Cossette, and J. Pepin, *Testing nursing students' clinical judgment in a patient deterioration simulation scenario: Development of a situation awareness instrument.* Nurse Education Today, 2016. **38**: p. 61-67.

40. Gantt, L.T., *Using the Clark Simulation Evaluation Rubric with associate degree and baccalaureate nursing students.* Nursing Education Perspectives, 2010. **31**(2): p. 101-105.

41. Benner, P., *From novice to expert: Excellence and power in clinical nursing*. 1984, Menlo Park, CA: Addison-Wesley.

42. Bloom, B.S., et al., *Taxonomy of educational objectives: The classification of educational goals*. 1956, New York: David McKay.

43. Furze, J., et al., *Clinical reasoning: Development of a grading rubric for student assessment.* Journal of Physical Therapy Education, 2015. **29**(3): p. 34-45.

44. Krathwohl, D.R., *A revision of Bloom's Taxonomy: An overview.* Theory Into Practice, 2002. **41**(4): p. 212-218.

45. O'Neill, E.S., N.M. Dluhy, and E. Chin, *Modelling novice clinical reasoning for a computerized decision support system.* Journal of Advanced Nursing, 2005. **49**(1): p. 68-77.

46. Bandura, A., *Self-efficacy: Toward a unifying theory of behavioral change.* Psychological Review, 1977. **84**(2): p. 191-215.

47. Bandura, A., *Social learning theory*. 1977, Englewood Cliffs, NJ: Prentice Hall.

48. White, K.A., *Development and validation of a tool to measure self-confidence and anxiety in nursing students during clinical decision making.* Journal of Nursing Education, 2014. **53**(1): p. 14-22.

49. Tanner, C.A., *Thinking like a nurse: A research-based model of clinical judgement in nursing.* Journal of Nursing Education, 2006. **45**(6): p. 204-211.

50. Pesut, D.J. and J. Herman, *Clinical reasoning: The art and science of critical and creative thinking*. 1999, Albany, NY: Delmar.

51. Levett-Jones, T., et al., *The 'five rights' of clinical reasoning: An educational model to enhance nursing students' ability to identify and manage clinically 'at risk' patients.* Nurse Education Today, 2010. **30**: p. 515-520.

52. Shin, H., C.G. Park, and K. Shim, *The Korean version of the Lasater Clinical Judgment Rubric: A validation study.* Nurse Education Today, 2015. **35**(1): p. 68-72.

53. Vreugdenhil, J. and B. Spek, *Development and validation of Dutch version of Lasater Clinical Judgment Rubric in hospital practice: An instrument design study.* Nurse Education Today, 2018. **62**: p. 43-51.

54. Georg, C., et al., *A rubric to assess students' clinical reasoning when encountering virtual patients.* Journal of Nursing Education, 2018. **57**(7): p. 408-415.

55. Shin, H., et al., *Validation of a new assessment tool for a pediatric nursing simulation module.* Journal of Nursing Education, 2014. **53**(11): p. 623-629.

56. Kim, S.J., et al., *Development of a simulation evaluation tool for assessing nursing students' clinical judgment in caring for children with dehydration.* Nurse Education Today, 2016. **37**: p. 45-52.

57. Liaw, S.Y., et al., *Development and psychometric testing of a Clinical Reasoning Evaluation Simulation Tool (CREST) for assessing nursing students' abilities to recognize and respond to clinical deterioration.* Nurse Education Today, 2018. **62**: p. 74-79.

58. Liou, S.R., et al., *The development and psychometric testing of a theory-based instrument to evaluate nurses' perception of clinical reasoning competence.* Journal of Advanced Nursing, 2016. **72**(3): p. 707-717.

59. Chatterjee, S., et al., *Assessing the surgical decision making abilities of novice and proficient urologists.* Journal of Urology, 2009. **181**(5): p. 2251-2256.

60. Derakhshandeh, Z., et al., *Psychometric characteristics of Clinical Reasoning Problems (CRPs) and its correlation with routine multiple choice question (MCQ) in Cardiology department.* Journal of Advances in Medical Education & Professionalism, 2018. **6**(1): p. 37-42.

61. Im, S., et al., *Assessing clinical reasoning abilities of medical students using clinical performance examination.* Korean Journal of Medical Education, 2016. **28**(1): p. 35-47.

62. Huwendiek, S., et al., *Electronic assessment of clinical reasoning in clerkships: A mixed-methods comparison of long-menu key-feature problems with context-rich single best answer questions.* Medical Teacher, 2017. **39**(5): p. 476-485.

63. Fida, M. and S.E. Kassab, *Do medical students' scores using different assessment instruments predict their scores in clinical reasoning using a computer-based simulation?* Advances in Medical Education & Practice, 2015. **6**: p. 135-41.

64. Beullens, J., E. Struyf, and B. Van Damme, *Do extended matching multiple-choice questions measure clinical reasoning?* Medical Education, 2005. **39**(4): p. 410-417.

65. Courteille, O., et al., *The use of a virtual patient case in an OSCE-based exam - A pilot study.* Medical Teacher, 2008. **30**(3): p. e66-e76.

66. Berger, A.J., et al., *Assessment of medical student clinical reasoning by "lay" vs physician raters: Inter-rater reliability using a scoring guide in a multidisciplinary objective structured clinical examination.* American Journal of Surgery, 2012. **203**(1): p. 81-86.

67. Tutticci, N., P.A. Lewis, and F. Coyer, *Measuring third year undergraduate nursing students' reflective thinking skills and critical reflection self-efficacy following high fidelity simulation: A pilot study.* Nurse Education in Practice, 2016. **18**: p. 52-59.

68. Smith, S., et al., *The development and preliminary validation of a rubric to assess medical students' written summary statements in virtual patient cases.* Academic Medicine, 2016. **91**(1): p. 94-100.

69. Fleiszer, D., et al., *Development and validation of a tool to evaluate the evolution of clinical reasoning in trauma using virtual patients.* Journal of Surgical Education, 2018. **75**(3): p. 779-786.

70. Adamson, K.A., *Rater bias in simulation performance assessment: Examining the effect of participant race/ethnicity.* Nursing Education Perspectives, 2016. **37**(2): p. 78-82.

71. Adamson, K.A. and S. Kardong-Edgren, *A method and resources for assessing the reliability of simulation evaluation instruments.* Nursing Education Perspectives, 2012. **33**(5): p. 334-9.

72. Ashcraft, A.S., et al., *Simulation evaluation using a modified Lasater Clinical Judgment Rubric.* Nursing Education Perspectives, 2013. **34**(2): p. 122-6.

73. Bussard, M.E., *Evaluation of clinical judgment in prelicensure nursing students.* Nurse Educator, 2018. **43**(2): p. 106-108.

74. Manetti, W., *Evaluating the clinical judgment of prelicensure nursing students in the clinical setting.* Nurse Educator, 2018. **43**(5): p. 272-276.

75. Strickland, H.P., M.H. Cheshire, and A.L. March, *Clinical judgment during simulation: A comparison of student and faculty scores.* Nursing Education Perspectives, 2017. **38**(2): p. 85-86.

76. Kautz, D., et al., *Building evidence for the development of clinical reasoning using a rating tool with the Outcome-Present State-Test (OPT) Model.* Southern Online Journal of Nursing Research, 2009. **9**(1): p. 8p-8p.

77. Amini, M., et al., *Psychometric characteristics of script concordance test (SCT) and its correlation with routine multiple choice question (MCQ) in internal medicine department.* Biomedical Research, 2017. **28**(19): p. 8397-8401.

78. Boulouffe, C., et al., *Assessing clinical reasoning using a script concordance test with electrocardiogram in an emergency medicine clerkship rotation.* Emergency Medicine Journal, 2014. **31**(4): p. 313-316.

79. Goos, M., et al., *Validation of undergraduate medical student script concordance test (SCT) scores on the clinical assessment of the acute abdomen.* BMC Surgery, 2016. **16**(1): p. 57.

80. Humbert, A.J., B. Besinger, and E.J. Miech, *Assessing clinical reasoning skills in scenarios of uncertainty: Convergent validity for a script concordance test in an emergency medicine clerkship and residency.* Academic Emergency Medicine, 2011. **18**(6): p. 627-634.

81. Kania, R.E., et al., *Online script concordance test for clinical reasoning assessment in otorhinolaryngology: The association between performance and clinical experience.* Archives of Otolaryngology - Head and Neck Surgery, 2011. **137**(8): p. 751-755.

82. Kazour, F., et al., *Using the Script Concordance Test to evaluate clinical reasoning skills in psychiatry.* Academic Psychiatry, 2017. **41**(1): p. 86-90.

83. Lambert, C., et al., *The Script Concordance Test in radiation oncology: Validation study of a new tool to assess clinical reasoning.* Radiation Oncology, 2009. **4**(7).

84. Ruiz, J.G., et al., *The Script Concordance Test as a measure of clinical reasoning skills in geriatric urinary incontinence.* Journal of the American Geriatrics Society, 2010. **58**(11): p. 2178-2184.

85. Sibert, L., et al., *Assessment of clinical reasoning competence in urology with the Script Concordance Test: An exploratory study across two sites from different countries.* European Urology, 2002. **41**(3): p. 227-233.

86. Sibert, L., et al., *On line clinical reasoning assessment with Script Concordance Test in urology: Results of a French pilot study.* BMC Medical Education, 2006. **6**(45).

87. Subra, J., et al., *Reliability and validity of the Script Concordance Test for postgraduate students of general practice.* European Journal of General Practice, 2017. **23**(1): p. 208-213.

88. Wan, M.S., E. Tor, and J.N. Hudson, *Improving the validity of script concordance testing by optimising and balancing items.* Medical Education, 2018. **52**(3): p. 336-346.

89. Groves, M., et al., *Analysing clinical reasoning characteristics using a combined methods approach.* BMC Medical Education, 2013. **13**: p. 144.

90. Gagnon, R., et al., *Script concordance testing: More cases or more questions?* Advances in Health Sciences Education, 2009. **14**(3): p. 367-375.

91. Dawson, T., et al., *Can script concordance testing be used in nursing education to accurately assess clinical reasoning skills?* Journal of Nursing Education, 2014. **53**(5): p. 281-286.

92. Funk, K.A., et al., *Experience with the script concordance test to develop clinical reasoning skills in pharmacy students.* Currents in Pharmacy Teaching and Learning, 2017. **9**(6): p. 1031-1041.

93. Cook, D.A., et al., *Management reasoning: Implications for health professions educators and a research agenda.* Academic Medicine, 2019. **94**(9): p. 1310-1316.

94. Boshuizen, H.P.A. and H.G. Schmidt, *The development of clinical reasoning expertise*, in *Clinical reasoning in the health professions*, J. Higgs, et al., Editors. 2019, Elsevier: Edinburgh. p. 57-66.

95. Roman-Cereto, M., et al., *Cultural adaptation and validation of the Lasater Clinical Judgment Rubric in nursing students in Spain.* Nurse Education Today, 2018. **64**: p. 71-78.

96. Lubarsky, S., et al., *The script concordance test: A new tool assessing clinical judgement in neurology.* Canadian Journal of Neurological Sciences, 2009. **36**(3): p. 326-331.