

ORIGINAL RESEARCH ARTICLE

Exploring the relationship of diagnostic reasoning, self-efficacy, and clinical reasoning of physical therapy students

Shala Cunningham^{1*}, Matthew Kessinger², Alejandra Mott³, Megan Carter⁴, Christine Plassmann⁵, Madison Downs⁶ and Catherine Blair South⁷

¹Department of Physical Therapy, Radford University, Roanoke, VA, USA; ²Virginia Commonwealth University – Tappahannock, Tappahannock, VA, USA; ³TheraSWIM, Loveland, CO, USA; ⁴Advanced Therapy Solutions, Greenville, SC, USA; ⁵Raleigh Court Rehabilitation, Roanoke, VA, USA; ⁶ATI Physical Therapy, Bolingbrook, IL, USA; ⁷AMN Healthcare, San Diego, CA, USA

Abstract

Purpose: One of the primary objectives of entry-level physical therapy education is to develop students' clinical reasoning (CR) skills to provide optimal, patient-centered care. However, CR is a highly contextualized concept, and the assessment of CR development to ensure students have the requisite skills for safe patient care prior to clinical experiences is challenging within an education program. Self-assessment tools may provide a time-efficient opportunity to assess students' CR development across their education. Both diagnostic reasoning and self-efficacy have been correlated with CR performance in physical therapy. This study aimed to explore the relationship between diagnostic reasoning, self-efficacy, and CR development among physical therapy students.

Methods: Diagnostic reasoning was assessed through the Diagnostic Thinking Inventory (DTI). Self-efficacy was measured by the New General Self-Efficacy (NGSE) scale and Physical Therapist Self-Efficacy (PTSE) scale. CR ability was evaluated through the Think Aloud Standardized Patient Examination (TASPE) performed during a standardized patient simulation, and scores for the CR performance criteria as assessed by clinical instructors on the Physical Therapist Clinical Performance Instrument Version 2006 (CPI).

Results: There was no correlation between self-assessment scores on the DTI, NGSE, PTSE, and CR performance assessed by faculty during a standardized patient simulation (TASPE) or clinical instructors using the midterm and final CPI during a 12-week full-time clinical experience.

Conclusion: This study was unable to identify a self-assessment tool or a student performance indicator that could accurately predict CR performance during upcoming full-time clinical experiences.

Keywords: *physical therapist education; clinical reasoning; self-efficacy*

Received: 1 August 2024; Revised: 14 January 2025; Accepted: 6 March 2025; Published: 17 April 2025

One key objective of entry-level physical therapy education programs is to assist students in the development and progression of clinical reasoning (CR) skills required to provide optimal, patient-centered care.¹⁻³ CR is developmental and is dependent on knowledge, skills, and accumulated experiences.^{3,4} It encompasses the cognitive processes, through which healthcare professionals assess patient information, formulate diagnoses, and develop tailored treatment plans.⁵ Within this framework, diagnostic reasoning plays a pivotal role, guiding clinicians in the systematic

analysis of patient data to arrive at accurate assessments and clinical decisions.^{5,6} Furthermore, CR is influenced by individual factors, including self-efficacy beliefs.⁷

As described by Huhn et al. in concept analysis, 'Clinical reasoning in physical therapy could be conceptualized as integrating cognitive, psychomotor and affective skills. It is contextual in nature and involves both therapist and client perspectives. It is adaptive, iterative and collaborative with the intended outcome being a biopsychosocial approach to patient/client management'.¹ Multiple

Poster presentations of the study were performed at APTA CSM 2023 and 2024.

models have been used to describe CR used by physical therapists (PTs), including the hypothetical deductive reasoning, pattern recognition, narrative reasoning, and dual process diagnostic reasoning, which encompasses the use of both hypothetical deductive reasoning and pattern recognition.^{2,8,9} Within these models are several advanced cognitive competencies, such as the application of fundamental scientific knowledge to clinical problems, targeted collection of data based on hypotheses, establishment of well-structured, problem-specific knowledge frameworks, synthesis of a comprehensive problem representation, and the utilization of metacognitive skills for the reflection of ongoing professional development.^{10–13} Fundamental for accurate patient assessment and subsequent treatment planning is diagnostic reasoning.¹⁴ Diagnostic reasoning is one component of CR involving the systematic gathering and analysis of patient data to formulate differential diagnoses and make informed clinical decisions.^{5,6}

Clinical reasoning is also influenced by individual factors. Self-efficacy plays a significant role in influencing how healthcare professionals approach and execute diagnostic and problem-solving processes.^{14,15} Self-efficacy, a central concept in social cognitive theory, refers to an individual's belief in their ability to perform specific tasks and achieve desired outcomes.^{16–18} Individuals with high self-efficacy are more motivated to engage in activities that improve their CR abilities, such as seeking feedback, pursuing additional training opportunities, and reflecting on their experiences. This proactive approach to learning contributes to ongoing improvement in CR proficiency.^{7,19} Higher levels of self-efficacy are associated with greater confidence in one's diagnostic skills, leading to more effective problem-solving and decision-making processes.^{16,20} Self-efficacy can also be categorized into general and specific self-efficacy.⁷ General self-efficacy is considered a stable construct reflective of cognition and personality characteristics. Specific self-efficacy is reflective of one's belief in the ability of successfully complete a task and can be related to an individual's professional capacity.⁷

Venskus and Craig propose that CR is characterized as 'developmental' due to its reliance on accumulated experience and the capacity to incorporate these experiences into clinical practice.⁷ Effective instructional strategies used across the health professions for developing CR skills include reflective thinking, metacognitive exercises, think aloud experiences, case-based learning, and simulation training.²¹ Simulation-based education is increasingly being used to replicate authentic patient encounters and meet objectives within physical therapist education programs. A scoping review on the use of simulation physical therapist education found that 47% of publications reported CR development as an objective of the simulation experience.²² Due to the contextual nature of CR, authentic patient experiences, such as simulation and clinical experiences, are considered ideal

for the observation and assessment of CR.^{1,11} To reliably determine CR performance, numerous observations of clinical performance are necessary.²³ However, this limits feasibility within an education program to assess the progression of CR skills of students based on the time commitment for longitudinal observations. To account for this limitation, many education programs for PT rely on clinical instructor feedback regarding CR performance during full-time clinical experiences using the Clinical Performance Instrument (CPI).²⁴

Despite the recognized relationship of both diagnostic reasoning and self-efficacy on CR development, limited research has investigated the interplay between these constructs in physical therapy students. Understanding the relationship between these constructs may provide instructors in physical therapy education programs an additional opportunity to recognize students who may benefit from supportive instruction to further develop their CR skills prior to full-time clinical experiences through the completion of self-perception questionnaires to measure individuals' beliefs and perceptions about their self-efficacy and diagnostic thinking abilities. This study aimed to explore the relationship between diagnostic reasoning, self-efficacy, and CR development among physical therapy students.

Methods

This study used a sample of convenience of 14 students in an entry-level Doctor of Physical Therapy program. The simulation experience was not considered a requirement of the curriculum and was optional for students participate in an authentic patient experience. Out of the 14 subjects, four students were in the first year, five from the second year, and five from the third year of the program. Eight students were female and 6 were male. The mean cumulative grade point average (GPA) at graduation was 3.80 with a range of 3.65 to 3.97. None of the students had direct patient access outside of the curriculum (i.e. rehabilitation technician, athletic trainer, and volunteer opportunities). All of the students were in good standing in the program.

The CR assessment was performed in the prior to upcoming clinical experience semesters in years 1, 2, and 3 of the program. Clinical experiences are typically 12 weeks in duration performed in the summer of year 1, summer of year 2, and spring of year 3. The first-year students had not yet completed a full-time clinical experience, the second-year students had completed one 12-week clinical experience, and the third-year students had completed one 6-week and one 12-week clinical experience. Students in the third year of the program completed a shortened clinical experience in the first year of the program due to COVID-19. See Fig. 1 for placement of the assessments. The students' options for clinical experience setting are limited by the curriculum to ensure students are competent in the skills for safe and effective patient

Year	Fall	Spring	Summer
0			Gross Anatomy
1	Exercise Physiology Kinesiology Clinical Medicine I (Integumentary) Patient Management I (Basic Exam Skills) Theory and Practice I (Patient Care Skills)	Neuroscience Scientific Inquiry I (Research Methods) Clinical Medicine II (Orthopedic Pathology) Patient Management II (Orthopedic Exam) Theory and Practice II (Exercise Instruction)	Year 1 Assessment Clinical Experience I (12 weeks)
2	Professional Affairs Differential Diagnosis and Imaging Psychosocial Elements of Illness and Disability Clinical Medicine III (Neurological Pathology) Neuromuscular Development and Control I	Health Policy and Administration Scientific Inquiry II (Statistics) Principles of Teaching and Learning Prosthetics and Orthotics Advanced Orthopedics Neuromuscular Development and Control II	Year 2 Assessment Clinical Experience II (12 weeks)
3	Preventative Health and Wellness Scientific Inquiry III (Capstone) Cardiopulmonary Patient Care Comprehensive Patient Care Pediatrics	Year 3 Assessment Research/Scientific Inquiry IV (Capstone Defense) Clinical Experience III (12 weeks)	

Fig. 1. Program organization and placement of clinical reasoning assessments.

care. Following completion of year one of the program, students may be placed in outpatient orthopedic clinic or skilled nursing facility for their full-time clinical experience. Following year 2, students may also participate in clinical settings for inpatient rehabilitation. In the third and final year of the program, students may choose a clinical experience across all settings and patient diagnoses.

Simulation experience

Students participated in a simulation experience with a standardized patient. The participants had the opportunity to review a short case summary before arriving at the clinical simulation center. The case review included behavioral objectives and a brief overview of clinical skills expected to be performed. In the simulation scenario, a patient with a complex medical and social history was evaluated in the home environment following hospitalization due to a recent hip fracture. The simulation scenario was reviewed by Certified Healthcare Simulation Educators at the simulation center to ensure the scenario aligned with the International Nursing Association for Clinical Simulation and Learning (INACSL) standards of best practice (INACSL).²⁵ The patient diagnosis of a hip fracture with open reduction and internal fixation was appropriate for application of the skills taught within the first year of the program. Students also learn how to perform a home assessment in the first semester of the education program. The patient's complex medical and social history, including polypharmacy, previous myocardial infarction, poorly controlled type II diabetes, macular degeneration, depression, and limited health literacy provided an opportunity to adapt the examination and treatment plan to meet the patient's unique health

and education needs. The total time of each session consisted of 100 min: 10 min for informed consent, 15 min for pre-briefing, 60 min for the simulation experience, and 15 for debriefing. The simulation experience was performed in a simulated apartment within a university-based simulation center. Digital video and audio recording of the scenario allowed the simulation to be live-streamed to observation stations for viewing by the faculty assessors and for future review of the simulation experience.

Outcome measurements

Prior to the simulation experience, the participants completed the NGSE, PTSE, and DTI as a form of self-reflection on their reasoning process. Each tool has been previously described as an assessment of student physical therapist CR. Two core faculty members in the Doctor of Physical Therapy program evaluated the student's performance during the simulation experience. Assessment commenced during the pre-briefing and extended through the debriefing process. The evaluation of student performance utilized the TASPE. In addition, following completion of the clinical experience, student scores on the CPI 2.0 were collected.

New general self-efficacy scale

The New General Self-Efficacy (NGSE) scale was developed and validated by Chen et al. in 2001.²⁶ It is a short, 8-item self-report questionnaire with a focus on ability to overcome challenges and accomplish goals. Items are scored on a five-point Likert scale with a minimum score of eight and a maximum score of 40. The initial psychometric evidence for this measure was based on the assessment of 316 undergraduates at a large mid-Atlantic

university. The internal consistency of the responses to these items ranges from 0.85 to 0.90. The scale demonstrates both content validity (88%) and predictive validity (comparative fit index, 0.90).²⁶ Additionally, the test–retest reliability was adequate ($r = 0.67$).²⁶ Although there are many measures of self-efficacy, the NGSE is often used as the standard for the development of tools for specific self-efficacy.²⁷

Physical therapist self-efficacy scale

The Physical Therapist Self-Efficacy (PTSE) scale is a five-item self-efficacy scale developed specific to CR in PT. Items are related to the examination, diagnosis, and treatment of patients including patient referral.⁷ Similar to the NGSE, the PTSE uses a five-point Likert scale with a minimum score of 5 and a maximum score of 25.⁷ The PTSE was validated for measuring self-efficacy in CR among PT students with scores increasing as students progressed through the 3-year Doctor of Physical Therapy (DPT) curriculum ($F = 22.134$; $df = 2$; $P = 0.000$).⁷ The PTSE has been shown to have good internal consistency (Cronbach's alpha 0.80).²⁸

Diagnostic thinking inventory

The Diagnostic Thinking Inventory (DTI) is a self-assessment tool developed by Bordage et al., which measures one's perception of flexibility in thinking (FT) and knowledge in memory.²⁹ The tool uses a knowledge-driven model of diagnostic thinking, which focuses on the organization and availability of medical knowledge that affects CR. The aim of the DTI is to differentiate CR skills between novice and expert clinicians. The tool consists of 41 questions: 21 questions on FT and 20 questions on structure in memory (SM). The items are scored on a six-point Likert scale with higher scores associated with greater CR ability. The maximum score for FT is 126, and the maximum score for SM is 120. The test and retest reliability for PT is positive ($r = 0.779$, $P < 0.001$).³⁰ The inventory had internal consistency with an overall alpha value of 0.846.³⁰ The DTI has been assessed for use with PT and differentiated between expert and novice PTs as well as licensed PTs and students.^{30–32}

Think aloud standardized patient examination

The Think Aloud Standardized Patient Examination (TASPE) was developed to assess students CR during a standardized patient simulation.¹¹ It utilizes a four-point scale ranging from excellent to poor during three think aloud sessions, in which the student reflects upon their CR for patient diagnosis and treatment planning across the hypothetical deductive reasoning process.¹¹ In this study, the first think aloud session was performed after reading the written case description of the patient. The second think aloud session occurred following the SP examination. Both the first and second think aloud sessions asked

the students to provide and justify the most likely diagnostic hypotheses based in the information gathered. The third think aloud session instructed students to describe and support the three most appropriate interventions based on the patient presentation.

Physical therapist clinical performance instrument version 2006

The Physical Therapist Clinical Performance Instrument (CPI) is the most commonly used instrument to assess Doctor of Physical Therapy (DPT) students' clinical competence.³³ It is a tool designed to evaluate PT student performance during their clinical experiences. The CPI is completed by both the clinical instructor and student at midterm and at the end of the clinical experience. It consists of 18 performance criteria. The rating scale is an ordered categorical scale consisting of six well-defined anchors.³³ A one on the scale corresponds to beginning performance for students and an 18 corresponds to beyond entry-level performance. The instrument demonstrated good validity for the CR performance criteria with the ability to distinguish student performance between the initial and final clinical experiences ($P < 0.001$). The CR performance criteria were utilized for this study.

Data analysis

Mean scores for each item on the assessment tools were determined. The Kruskal Wallis was utilized to assess the ability of the tools to distinguish student performance on CR based on clinical experience and progression in the program. Furthermore, Spearman's rho calculation explored the relationship between performance on the NGSE, PTSE, DTI, TASPE, and scoring on the CPI performance criteria for CR as determined at midterm (6 weeks) and completion (12 weeks) during a full-time clinical experience.

Results

Second year students demonstrated the highest overall composite scores on the DTI. Third-year students scored lower on 20 individual items within the DTI compared to second-year students. Notably, second-year students scored lower than first-year students for the composite score for items within SM with third year students scoring highest. Student's overall performance on the DTI can be found in Table 1. The mean scores for each item can be found in Appendix A.

Although there was a slight improvement in means of the TASPE between the first- and second-year students, there appeared to be a plateau on the majority of think aloud sections between second- and third-year students, suggesting similar considerations for the justification of the hypotheses following the patient examination and the proposed treatment between the two cohorts. Student scores on the TASPE can be seen in Table 2.

Table 1. DTI section scores and Kruskal Wallis comparison

DTI section scores	Mean first year	Mean second year	Mean third year	Kruskal Wallis (sig.)
DTI Flexibility in Thinking Composite Score	105.00	113.25	107.00	0.907
DTI Structure in Memory Composite Score	27.75	26.50	30.60	0.063
DTI Total Composite Score	132.75	139.75	137.60	0.372

DTI, Diagnostic Thinking Inventory; sig., statistical significance.

Table 2. Student scores on the Think Aloud Standardized Patient Examination by cohort and Kruskal Wallis comparison

Question TASPE	Mean first year	Mean second year	Mean third year	Kruskal Wallis (sig.)
Provide and justify the three most likely diagnostic hypotheses based on positive and negative findings found in the written description of the patient				
Diagnostic Hypothesis #1	1.75	1.80	1.80	0.988
Diagnostic Hypothesis #2	1.25	1.60	1.60	0.813
Diagnostic Hypothesis #3	1.25	1.60	1.60	0.839
Composite Score Think Aloud 1	4.75	5.00	5.00	0.930
Provide and justify the three most likely diagnostic hypotheses based on positive and negative findings found in the written description of the patient and the physical examination				
Diagnostic Hypothesis #1	1.25	1.60	1.80	0.553
Diagnostic Hypothesis #2	1.25	1.60	1.80	0.553
Diagnostic Hypothesis #3	1.00	1.40	1.40	0.627
Composite Score Think Aloud 2	3.50	4.60	5.00	0.631
Provide and justify the three most appropriate interventions for the clinical presentation with your immediately preceding first diagnostic hypothesis considered				
Diagnostic Hypothesis #1	1.33	1.80	1.80	0.713
Diagnostic Hypothesis #2	0.67	1.40	1.60	0.426
Diagnostic Hypothesis #3	0.0	1.60	1.80	0.091
Composite Score Think Aloud 3	2.0	4.80	5.20	0.080

TASPE, Think Aloud Standardized Patient Examination; sig., significance.

Both second- and third-year students scored higher on the PTSE and NGSE than first-year students. However, second-year students scored higher than third-year students on the majority of questions. See Table 3 for the PTSE and NGSE mean scores by cohort. Students did demonstrate a statistically significant improvement in scores on the CPI as they progressed in the program when exploring the Kruskal Wallis. On post hoc analysis, this difference was seen between the first- and second-year students. See Table 4 for CPI scores. However, there was no correlation between diagnostic reasoning (DTI), self-efficacy (PTSE and NGSE), and performance of CR as assessed within the program (TASPE) and on clinical experiences (CPI). The correlations can be found in Table 5.

Discussion

The results of this study present a departure from prior research findings regarding the relationship between diagnostic reasoning (DTI), self-efficacy (NGSE and PTSE), and CR ability among physical therapy students.^{26,30–32} Contrary to expectations based on existing literature, our

analysis did not reveal a significant correlation between DTI scores and CR ability, nor between self-efficacy and CR ability. Furthermore, student's performance during a standardized patient simulation (TASPE) in this study did not correlate to CR performance during the subsequent full-time clinical experience based on CPI scores.

One possible explanation for this discrepancy could be the context of our study including students across all 3 years of a physical therapy program. Although the DTI has been shown to discriminate between student, novice, and expert clinicians, the DTI does not have normative ranges to describe the adequate level of reasoning for students. It has been suggested that the tool may not discriminate small gains in diagnostic reasoning skills.³⁴ Furthermore, self-efficacy is believed to be task specific, which suggests that participants level of self-efficacy in physical therapy practice (PTSE) may not necessarily correlate with the perceived level of general self-efficacy (NGSE), which is a personal trait that remains relatively stable across situations.¹⁷ It was expected that self-efficacy should develop over time as students acquire additional knowledge and experience

Table 3. Cohort mean scores on the Physical Therapist Self-Efficacy and New General Self-Efficacy and Kruskal Wallis comparison

PTSE	Mean first year	Mean second year	Mean third year	Kruskal Wallis (sig.)
I am confident that I know when to perform specific tests for physical therapist practice.	4.0	4.0	3.8	1.000
I will know when it is time to refer a patient to another practitioner.	4.0	4.0	4.4	0.561
In a general physical therapy context, I am confident that I wouldn't not miss primary medical disease.	2.5	3.75	3.6	0.131
I believe that I can manage general physical therapy problems.	4.0	4.5	4.4	0.507
In a general physical therapy context, when facing a difficult case, I am certain that I can make the right management decisions.	2.5	3.75	3.4	0.201
NGSE	Mean first year	Mean second year	Mean third year	Kruskal Wallis (sig.)
I will be able to achieve most of the goals that I have set for myself.	4.0	4.75	4.4	0.241
When facing difficult tasks, I am certain that I will accomplish them.	4.0	4.25	3.8	0.325
In general, I think that I can obtain outcomes that are important to me.	4.0	4.75	4.4	0.241
I believe that I can succeed at most any endeavor to which I set my mind	4.0	4.75	4.8	0.142
I will be able to successfully overcome many challenges.	4.0	4.5	4.6	0.380
I am confident that I can perform effectively on many different tasks.	4.0	4.5	4.2	0.608
Compared to other people, I can do most tasks very well.	3.5	4.0	3.4	0.349
Even when things are tough, I can perform quite well.	3.5	4.25	3.6	0.433

PTSE, Physical Therapist Self-Efficacy Scale; NGSE, New General Self-Efficacy Scale; sig., significance.

Table 4. Clinical performance instrument clinical reasoning

Performance criteria clinical reasoning	Mean first year	Mean second year	Mean third year	Kruskal Wallis sig.	Year 1–2 Mann U Whitney sig.	Year 2–3 Mann U Whitney sig.
Midterm Mean	3.75	10.2	13.4	0.011	0.016	0.095
Final Mean	5.75	15.6	16.6	0.014	0.036	0.548

sig., significance

in practice.¹⁸ However, similar to results noted by Venskus and Craig, second-year physical therapy students in this study demonstrated higher self-efficacy on several items in the NGSE and PTSE compared to third-year students.⁷ Similarly, the third-year students' preparation for the transition from being a student to practicing professional may have contributed to the decrease in scores.⁷

Student performance on the TASPE during the standardized patient simulation did not correlate to performance in the clinic. Several factors may have influenced this result. Two faculty members completed the TASPE. Although previous training on the tool had occurred, and inter-rater reliability was established for scoring the tool prior to the study, variability in expectations for performance may have been present. The practice setting for the simulation experience was a home health environment. This setting was chosen because students who had completed a full-time clinical experience had not yet had the opportunity to participate in home health care. All students focused on patient safety in the home as a key priority of the examination and treatment, which may have limited the ability to distinguish CR

development using the TASPE. The cues within the TASPE for reflection focused on the three most likely hypotheses and may not have been sufficient to encourage students to discuss how the patient's complex medical history influenced their clinical decisions. The simulation was limited to a single patient encounter. Follow-up sessions with the patient, once safety in the home was established, may have allowed students more opportunity to individualize reassessment and treatment procedures to match the patient's complex presentation.

Finally, the second-year students performed at a high level (entry-level practice) for the CR performance criteria during their clinical experience. This may be attributed to the practice setting for the students. Of the five second-year students, two students completed their clinical experience in an outpatient setting, two were assigned to a skilled nursing facility, and one practiced in an acute care setting treating mainly post-operative orthopedic patients. However, it is difficult to determine complexity of patients based on practice setting. Of the five clinical instructors, three had training through the American Physical

Table 5. Spearman's Rho correlations New General Self-Efficacy, Physical Therapist Self-Efficacy, Diagnostic Thinking Inventory, Think Aloud Standardized Patient Examination, and Clinical Performance Instrument

Tool		NGSE	PTSE	DTI composite	DTI FT	DTI SM	TASPE composite	CPI midterm	CPI final
NGSE	rs	1.000	0.861	0.271	0.214	0.193	0.298	0.031	0.403
	sig.		<0.001*	0.416	0.527	0.570	0.374	0.929	0.219
PTSE	rs	0.861	1.000	0.375	0.391	0.118	0.195	0.021	0.347
	sig.	<0.001*		0.256	0.235	0.729	0.565	0.951	0.296
DTI composite	rs	0.273	0.375	1.000	0.766	0.472	0.039	0.076	0.135
	sig.	0.416	0.256		0.002*	0.103	0.899	0.805	0.660
DTI FT	rs	0.214	0.391	0.214	1.00	0.206	0.068	0.137	0.047
	sig.	0.527	0.235	0.527		0.500	0.826	0.656	0.884
DTI SM	rs	0.193	0.118	0.193	0.206	1.00	0.313	0.318	0.363
	sig.	0.579	0.729	0.570	0.500		0.298	0.290	0.239
TASPE composite	rs	0.298	0.195	0.039*	0.068	0.313	1.000	0.169	0.381
	sig.	0.374	0.565	0.899	0.826	0.298		0.563	0.179
CPI midterm	rs	0.031	0.021	0.076	0.137	0.318	0.169	1.00	0.896
	sig.	0.929	0.951	0.805	0.656	0.290	0.563		<0.001*
CPI final	rs	0.403	0.347	0.135	0.047	0.363	0.179	0.896	1.00
	sig.	0.219	0.296	0.660	0.884	0.239	0.896	<0.001*	

rs, Spearman's Rho; sig., significance; NGSE, New General Self-Efficacy Scale; PTSE, Physical Therapist Self-Efficacy Scale; DTI, Diagnostic Thinking Inventory; DTI FT, Diagnostic Thinking Inventory Flexibility in Thinking; DTI SM, Diagnostic Thinking Inventory Structure in Memory; TASPE, Think Aloud Standardized Patient Examination; CPI, Clinical Performance Instrument.

* = statistically significant.

Therapy Association completing the Credentialed Clinical Instructor Program Level 1. All clinical instructors receive support from the program's Director of Clinical Education for the scoring of the CPI.

Limitations

There were multiple limitations for this pilot study. This study utilized a small sample of convenience from a single physical therapy education program limiting the generalizability of results. Students in the third year of the program completed a shortened clinical experience (6 weeks) in the first year of the program due to COVID-19 restrictions. An additional 6 weeks of experience in the clinic may affect specific self-efficacy perceptions. However, all students met the program expectation for the first and second clinical experience.

The simulation experience was designed to allow students across the program to be able to perform an examination and develop a treatment plan. However, the context of the scenario being in a home setting may have inadvertently resulted in similar reasoning focused on patient safety across the cohorts. The home setting also did not represent the clinical environment that the students were entering. The change in context may have affected the ability of the student performance in simulation to predict performance in the clinic.

Although two faculty scored each student independently on the TASPE, bias may have been introduced as the faculty were aware of cohort the students were assigned. Similar bias may have been present for clinical instructors when scoring the CPI. Considerations

for assessing student CR performance may have been inconsistent between core faculty and clinical instructors based on preconceived expectations for performance and an understanding of what constitutes entry-level performance. The clinical instructors did not receive additional training from the study team for scoring CR performance on CPI, which could have resulted in inconsistency in scoring the item. Furthermore, the test-retest reliability of the PTSE is unknown and may have influenced results.

Conclusion

The findings of this study challenge previous research regarding the relationship between diagnostic reasoning, self-efficacy, and CR ability among physical therapy students. Contrary to expectations, the self-assessment tools (DTI, NGSE, and PTSE) and student performance during a standardized patient simulation (TASPE) were unable to predict CR performance during upcoming full-time clinical experiences. Future research should explore alternative simulation scenarios that match the setting of the upcoming clinical experience and assessments that may better capture the intricacies of CR in physical therapy education and the relationship to performance in clinical practice.

Conflict of interests and funding

The authors have no conflicts of interest.

Ethical statement

This study was approved by the Radford University Institutional Review Board (2021-411-RUC).

References

- Huhn K, Gilliland S, Black L, Wainwright S, Christensen N. Clinical reasoning in physical therapy: a concept analysis. *Phys Ther* (2019) 99(4): 440–56. doi: 10.1093/ptj/pzy148
- Prokop T. Use of the dual-processing theory to develop expert clinical reasoning in physical therapy students. *J Phys Ther Educ* (2018) 32(4): 355–9. doi: 10.1097/JTE.0000000000000062
- Gilliland S, Wainwright SF. Patterns of clinical reasoning in physical therapist students. *Phys Ther* (2017) 97: 499–511. doi: 10.1093/ptj/pzx028
- Noll E, Key A, Jensen G. Clinical reasoning of an experienced physiotherapist: insight into clinician decision-making regarding low back pain. *Physiother Res Int* (2001) 6(1): 40–51. doi: 10.1002/pri.212
- Rajkomar A, Dhaliwal G. Improving diagnostic reasoning to improve patient safety. *Permanente J* (2011) 15(3): 68–72. doi: 10.7812/TPP/11-098
- Nordick CL. Integrating strategies for improving diagnostic reasoning and error reduction. *J Am Assoc Nurse Pract* (2021) 33(5): 366–72. doi: 10.1097/JXX.0000000000000464
- Venskus DG, Craig JA. Development and validation of a self-efficacy scale for clinical reasoning in physical therapists. *J Phys Ther Educ* (2017) 31: 14–20. doi: 10.1097/00001416-201731010-00005
- Elstein AS, Schwarz A. Clinical problem solving and diagnostic decision making: a selective review of the literature. *BMJ* (2002) 324: 729–32. doi: 10.1136/bmj.324.7339.729
- Jensen G, Shepard KF, Gwyer J, et al. Attribute dimensions that distinguish master and novice physical therapy clinicians in orthopedic settings. *Phys Ther* (1992) 72(10): 30–41. doi: 10.1093/ptj/72.10.711
- Castillo JM, Park YS, Harris I, et al. A critical narrative review of transfer of basic science knowledge in health professions education. *Med Educ* (2018) 52(6): 592–604. doi: 10.1111/medu.13519
- Fu W. Development of an innovative tool to assess student physical therapists' clinical reasoning competency. *J Phys Ther Educ* (2015) 29(4): 14–26. doi: 10.1097/00001416-201529040-00004
- Lambe KA, O'Reilly G, Kelly BD, et al. Dual-process cognitive interventions to enhance diagnostic reasoning: a systematic review. *BMJ Qual Saf* (2016) 25: 808–20. doi: 10.1136/bmjqs-2015-004417
- Olson A, Rencic J, Cosby K, et al. Competencies for improving diagnosis: an interprofessional framework for education and training in health care. *Diagnosis* (2019) 6(4): 335–41. doi: 10.1515/dx-2018-0107
- Klassen RM, Klassen LRL. Self-efficacy beliefs of medical students: a critical review. *Perspect Med Educ* (2018) 7(2): 76–82. doi: 10.1007/S40037-018-0411-3
- Abusubiah M, Walshe N, Creedon R, et al. Self-efficacy in the context of nursing education and transition to practice as a registered practitioner: a systematic review. *Nurs Open* (2023) 10(10): 6650–67. doi: 10.1002/nop2.1931
- Flowers M, Yates C, Fletcher J, et al. Does dosing of pediatric experiential learning impact the development of clinical reasoning, self-efficacy, and critical thinking in DPT students? *J Allied Health* (2020) 49: 190–6.
- Stajkovic AD, Luthans F. Social cognitive theory and self-efficacy: implications for motivation theory and practice. In: Steers RM, Porter LW, Bigley GA, eds. *Initial considerations*. New York: McGraw-Hill/Irwin; 2002, pp. 126–40.
- Bandura A. Self-efficacy: toward a unifying theory of behavioral change. *Psychol Rev* (1977) 84: 191–215. doi: 10.1037//0033-295X.84.2.191
- Mavis B. Self-efficacy and OSCE performance among second year medical students. *Adv Health Sci Educ Theory Pract* (2001) 6(2): 93–102.
- Lowe L, Flowers M, Fletcher J, et al. The effects of experiential learning on perceived clinical reasoning and self-efficacy of doctor of physical therapy students. *J Phys Ther Educ* (2022) 36(2): 122–7. doi: 10.1097/JTE.0000000000000225
- Higgs J, Jensen G, Loftus S, et al. *Clinical reasoning in the health professions*. 4th ed. Edinburgh: Elsevier; 2019.
- Stockert B, Silberman N, Rucker J, et al. Simulation-based education in physical therapist professional education: a scoping review. *Phys Ther* (2022) 102(12): 133. doi: 10.1093/ptj/pzac133
- Rencic J, Schuwirth, Lambert WT, et al. Clinical reasoning performance assessment: using situated cognition theory as a conceptual framework. *Diagnosis* (2020) 7(3): 241–9. doi: 10.1515/dx-2019-0051
- Huhn K, Black L, Christensen N, et al. Clinical reasoning: survey of teaching methods, integration, and assessment in entry-level physical therapist academic education. *J Phys Ther Educ* (2017) 97: 175–86. doi: 10.2522/ptj.20150320
- Watts PI, Rossler K, Bowler F, et al. Onward and upward: introducing the healthcare simulation standards of best practice™. *Clin Simul Nurs* (2021) 58: 1–4. doi: 10.1016/j.ecns.2021.08.006
- Chen G, Gully SM, Eden D. Validation of a new general self-efficacy scale. *Organ Res Methods* (2001) 4: 62–83. doi: 10.1177/109442810141004
- Scherbaum CA, Cohen-Charash Y, Kern MJ. Measuring general self-efficacy: a comparison of three measures using item response theory. *Educ Psych Measure* (2006) 66(6): 1047–63. doi: 10.1177/0013164406288171
- Campbell D, Brismée JM, Allen B, et al. Doctor of physical therapy students' clinical reasoning readiness and confidence treating with telehealth: a United States survey. *J Clin Educ Phys Ther* (2022) 4: 1–10. doi: 10.52214/jcept.v4.8537
- Bordage G, Grant J, Marsden P. Quantitative assessment of diagnostic ability. *Med Educ* (1990) 24(5): 413–25. doi: 10.1111/j.1365-2923.1990.tb02650.x
- Hamzeh H, Madi M. Using the diagnostic thinking inventory in musculoskeletal physiotherapy: a validity and reliability study. *Physiother Res Int* (2021) 26(2): e1895. doi: 10.1002/pri.1895
- Jones U. The reliability and validity of the Bordage, Grant & Marsden diagnostic thinking inventory for use with physiotherapists. *Med Teach* (1997) 19: 133–40. doi: 10.3109/01421599709019366
- Salles FL, Ferreira DM, Lima FF, et al. Evaluation of the reliability of diagnostic thinking inventory for physical therapist and of the development of clinical reasoning in physiotherapy students. *Eur J Physiother* (2020) 24: 208–18. doi: 10.1080/21679169.2020.1842494
- Roach KE, Frost JS, Francis NJ, et al. Validation of the revised physical therapist clinical performance instrument (PT CPI): version 2006. *Phys Ther* (2012) 92(3): 416–28. doi: 10.2522/ptj.20110129
- Harris KI, Rowat JS, Suneja M. Embedding a longitudinal diagnostic reasoning curriculum in a residency program using a bolus/booster approach. *Diagnosis* (2020) 7: 21–25. doi: 10.1515/dx-2019-0023

***Shala Cunningham**

Radford University

101 Elm Ave SE

Roanoke, VA 24013, USA

Email: scunningham4@radford.edu

Appendix A. Mean DTI scores for all three cohorts and Kruskal Wallis comparison

DTI Question and Anchors Flexibility in Thinking questions 1–32 Structure in Memory questions 33–41	Mean First Year	Mean Second Year	Mean Third Year	Kruskal Wallis (sig)
1. When the patient presents symptoms, (1) <i>I think of the symptoms in the precise words used by the patient</i> , (5) <i>I think of the symptoms in more abstract terms than the expressions actually used</i> (e.g. '4-day duration' becomes 'acute'; 'two-hands' becomes 'bilateral')	2.25	4.75	4.20	0.35
2. In considering each diagnosis, (1) <i>I try to evaluate their relative importance</i> , (5) <i>I try to give them equal importance or weighting</i>	1.50	3.25	2.00	.383
3. In thinking of diagnostic possibilities, (1) <i>I think of diagnostic possibilities early on in the case</i> , (5) <i>First I collect the clinical information and then I think about it</i>	3.00	2.0	3.00	.620
4. When I am interviewing a patient, (1) <i>I often seem to get one idea stuck in my mind about what might be wrong</i> , (5) <i>I usually find it easy to explore various possible diagnosis</i>	4.00	3.5	3.00	.568
5. Throughout the interview, (1) <i>If I follow the patient's line of thought, I tend to lose my own thread</i> , (5) <i>I can still keep my own ideas clear even if I follow the patient's line of thought</i>	3.00	4.0	3.80	.632
6. When it comes to making up my mind about a diagnosis, (1) <i>I do not mind postponing my diagnostic decisions about a case</i> , (5) <i>I feel obliged to go for one diagnosis or another even if I am not very certain</i>	2.50	2.5	2.40	.955
7. Once the patient has clearly presented his symptoms and signs, (1) <i>I think about them in my mind in the patient's own words</i> , (5) <i>I translate them in my mind into medical terms</i> (e.g. 'numbness' becomes 'paraesthesia' or 'paralysis')	3.50	4.5	4.20	.689
8. In relation to the routine history, (1) <i>I often feel that I did not sufficiently cover the routine history</i> , (5) <i>I usually cover the routine history to my satisfaction</i>	3.50	5.75	4.20	.063
9. As the patient tells his story and the case unfolds, (1) <i>I often find it difficult to remember what has been said</i> , (5) <i>I can usually keep track in my mind of what has been said</i>	4.00	5.25	4.40	.400
10. During the course of the interview, I find that, (1) <i>Some key pieces of information seem to leap out at me</i> , (5) <i>It is often difficult to know which items of information to latch on to</i>	1.75	2.25	2.00	.687
11. When I cannot make sense of the patient's symptoms, (1) <i>I move on and gather new information to trigger new ideas</i> , (5) <i>I ask the patient to define those symptoms more clearly</i>	3.50	4.75	4.00	.359
12. In considering diagnostic possibilities, (1) <i>I often come up with unlikely diagnoses</i> , (5) <i>I am usually in the right area</i>	4.00	4.5	4.80	.364
13. While I am collecting information about a patient, (1) <i>The various items of information usually seem to group themselves together in my mind</i> , (5) <i>I often have difficulty seeing how the pieces of information relate to each other</i>	2.75	2.25	3.20	.232
14. When the diagnosis becomes known and I realise that I have missed it initially, (1) <i>It is often because I knew the disease but failed to think about it</i> , (5) <i>It is often because I did not know enough about the disease</i>	4.50	2.75	3.60	.117
15. During the clinical interview, (1) <i>I cannot bring myself to dismiss some information as irrelevant</i> , (5) <i>I am quite happy to dismiss some information as irrelevant</i>	4.00	4.25	4.80	.320
16. When I cannot make sense of the patient's symptoms and signs, (1) <i>I move on to get new information and a new perspective</i> , (5) <i>I look at them from a different perspective before moving on</i>	4.25	3.5	3.60	.672
17. When I consider a number of possible diagnoses, (1) <i>The diagnoses tend to be related to one another</i> , (5) <i>The diagnoses tend to be scattered</i>	2.00	2.25	2.20	.907
18. When a possible diagnosis comes to my mind, (1) <i>I usually find myself anticipating possible abnormal signs and symptoms that go with that diagnosis</i> , (5) <i>Quite often, it does not help me to decide what to ask the patient next</i>	2.00	2.25	2.40	.728
19. When I know very little about a particular type of disease, (1) <i>I can still usually come up with a diagnosis</i> , (5) <i>I have great difficulty in reaching a diagnosis</i>	4.25	3.5	4.00	.472
20. In considering the patient's signs and symptoms, (1) <i>I think about each in absolute terms as stated by the patient</i> , (5) <i>I think of them in terms of possible opposites</i> (e.g. <i>progressive vs. sudden; unilateral vs. bilateral; spastic vs. flaccid</i>)	4.00	4.25	3.60	.698
21. When I know a lot about a particular type of disease and have to make a diagnosis, (1) <i>I find it relatively easy to pin down a diagnosis</i> , (5) <i>I often seem to be all over the place and have difficulty pinning down a diagnosis</i>	2.75	3.25	2.20	.405
22. As the history progresses and I already have some ideas about the possible diagnosis(es), (1) <i>New information often makes me have more ideas</i> , (5) <i>New information does not often make me have more ideas</i>	1.75	2.75	2.60	.251

(Continued)

Appendix A. (Continued) Mean DTI scores for all three cohorts and Kruskal Wallis comparison

DTI Question and Anchors	Mean First Year	Mean Second Year	Mean Third Year	Kruskal Wallis (sig)
Flexibility in Thinking questions 1–32				
Structure in Memory questions 33–41				
23. When I am taking a history, I find that, (1) <i>I can get new ideas just by going over the existing information in my mind</i> , (5) <i>I need to have new information to make me have a new idea about the case</i>	3.5	2.75	2.60	.837
24. When patients use imprecise or ambiguous expressions, (1) <i>I let them go on to maintain the flow of the interview</i> , (5) <i>I make them clarify precisely what they mean before going on</i>	4.25	3.5	3.60	.687
25. After an interview with a patient, (1) <i>I rarely think of other things that I should have asked in relation to the patient's disorder</i> , (5) <i>I often think of other things that that I should have asked in relation to the patient's disorder</i>	5.25	5.25	5.40	.875
26. When a piece of information comes along and makes me think of a possible diagnosis, (1) <i>It often makes me go back to previous information to see if things fit together or not</i> , (5) <i>It rarely makes me review the information that I have gathered previously</i>	1.5	2.75	1.40	.062
27. In relation to the diagnosis I eventually make, (1) <i>I usually have very few doubts</i> , (5) <i>I often feel too uncertain for my own comfort</i>	3.25	2.75	3.00	.854
28. In making a diagnostic decision, (1) <i>I decide by considering each possible diagnosis separately on its own merits</i> , (5) <i>I decide by comparing and contrasting the various possible diagnosis</i>	4.75	3.75	4.60	.482
29. When I know a lot about a particular type of disease and have to make a diagnosis, (1) <i>I check up on most possibilities before reaching a decision</i> , (5) <i>I often have lots of ideas that I don't explore further</i>	2.50	4.0	2.00	.177
30. As the case unfolds, (1) <i>I do not find it useful to summarise as I go along</i> , (5) <i>I periodically take stock of the data and my ideas</i>	5.00	3.5	3.80	.257
31. When I reach my diagnostic decisions, (1) <i>There is often left-over information I have just forgotten about</i> , (5) <i>I usually will have considered all the information</i>	4.00	4.25	3.40	.536
Flexibility in Thinking Composite Score	105.00	113.25	107.00	.907
32. When I have got an idea about what might be wrong with the patient, (1) <i>I feel most comfortable if I can follow it up without being diverted</i> , (5) <i>I feel happy to go off on another tack and come back to my original ideas later</i>	2.25	2.75	3.00	.806
33. When I come up with a broad ideas as to what might be wrong with the patient, (1) <i>I can usually proceed to a specific diagnosis</i> , (5) <i>I find it difficult to put it into specific terms</i>	3.50	3.00	3.40	.808
34. Throughout the interview, (1) <i>I manage to test my ideas even if I let the patient control the interview</i> , (5) <i>I am only successful if I can control the direction of the interview</i>	2.00	2.25	3.20	.256
35. In relation to choosing from among the diagnostic ideas that I have, (1) <i>I am usually not capable of wholly ruling out any of the ideas I have had</i> , (5) <i>I am capable of ruling out most of my ideas completely</i>	4.00	4.25	3.80	.612
36. Once I have made up my mind about a patient, (1) <i>I am prepared to change my mind</i> , (5) <i>I really do not like to change my mind</i>	2.25	3.00	2.20	.506
37. When I consider my diagnostic ideas, I do so on the basis of, (1) <i>The case as a whole so far</i> , (5) <i>A few outstanding symptoms or signs</i>	2.75	2.50	2.60	.942
38. If I do not know what to make of a clinical interview, (1) <i>I can readily see the information in new ways</i> , (5) <i>I find it difficult to see the information in new ways</i>	3.25	3.50	4.00	.724
39. When I order laboratory tests, (1) <i>I do it as part of the clinical routine investigation</i> , (5) <i>I consider each diagnosis separately on its own merits</i>	5.00	4.00	5.00	.667
40. In considering diagnostic possibilities, (1) <i>I compare and contrast the possible diagnoses</i> , (5) <i>I consider each diagnosis separately on its own merits</i>	2.50	2.25	2.60	.942
41. In terms of the way I conduct an interview, (1) <i>I usually cover the ground that I need to during the interview</i> , (5) <i>Quite often I do not ask all the questions that I should at the time</i>	2.50	1.75	3.80	.025
Structure in Memory Composite Score	27.75	26.50	30.60	.063
DTI Diagnostic Thinking Inventory				
Sig Statistical Significance				