

Development and Validation of the Systems Thinking Scale

Mary A. Dolansky, PhD, RN¹, Shirley M. Moore, PhD, RN¹, Patrick A. Palmieri, PhD², and Mamta K. Singh, MD, MS^{3,4}



¹Frances Payne Bolton School of Nursing, Case Western Reserve University, 10900 Euclid Ave., Cleveland, OH, USA; ²Traumatic Stress Center, Department of Psychiatry, Summa Health System, Akron, OH, USA; ³Office of Curricular Affairs, Case Western Reserve University School of Medicine, Cleveland, OH, USA; ⁴Coordinating Center, Centers of Excellence in Primary Care Education, Office of Academic Affiliations; Veterans Affairs, Cleveland, USA.

PURPOSE: Systems thinking is the ability to recognize and synthesize patterns, interactions, and interdependencies in a set of activities and is a key component in quality and safety. A measure of systems thinking is needed to advance our understanding of the mechanisms that contribute to improvement efforts. The purpose of this study was to develop and conduct psychometric testing of a systems thinking scale (STS).

METHODS: The development of the STS included obtaining national quality and safety experts' conceptual domains of systems thinking and the generation of a provisional set of items. Further psychometric analyses were conducted with interprofessional healthcare faculty ($N=342$) and students ($N=224$) engaged in quality improvement initiatives and education.

RESULTS: Of the 26 items identified in the development phase, factor analyses indicated three factors: (1) system thinking (20 items), (2) personal effort (2 items), and (3) reliance on authority (4 items). The six items from factors 2 and 3 were omitted due to low factor loadings. Test-retest reliability of the 20-item STS was performed on 36 healthcare professionals and a correlation of 0.74 was found. Internal consistency testing on a sample of 342 healthcare professionals using Cronbach's alpha showed a coefficient of 0.89. Discriminant validity was confirmed with three groups of healthcare professions students ($N=102$) who received high, low, or no dose levels of systems thinking education in the context of process improvement.

CONCLUSIONS: The 20-item STS is a valid and reliable instrument that is easy to administer and takes less than 10 min to complete. Further research using the STS has the potential to advance the science and education of quality improvement in two main ways: (1) increase understanding of a critical mechanism by which quality improvement processes achieve results, and (2) evaluate the effectiveness of our education to improve systems thinking.

INTRODUCTION

The ability to engage in systems thinking is viewed as a key component in the success of quality improvement initiatives and is critical to systems-based practice (SBP), one of the six Accreditation Council of Graduate Medical Education (ACGME)¹ core competencies for physicians in training and a core competence in the Future of Nursing Report.² Systems thinking is defined as the ability to recognize and synthesize patterns, interactions, and interdependencies in a set of activities designed for a specific purpose.³ This includes the ability to recognize patterns in the interactions and an understanding of how actions can reinforce or counteract each other.³ International health educators have recommended systems thinking as a necessary part of health education and research training.^{4, 5} Additionally, several initiatives point to conceptual system-based models designed to guide healthcare providers to improve systems thinking resulting in patient safety and quality such as the (a) London protocol,⁶ (b) systems engineering initiative for patient safety,⁷ (c) significant event analysis,⁸ and (d) systems awareness model that are applicable for both education and practice.⁹

Recently, many medical schools have added health systems science to their curricula as a concept and set of skills that are foundational to medicine. Health Systems Science is defined as "the principal, methods, and practice of improving quality, outcomes, and costs of health delivery for patient and population within systems of medical care. It helps us to see how to change systems more effectively and to act more in tune with the larger processes of the natural and economic world."¹⁰ Systems thinking is considered an "interlinking domain" in health systems science and an essential skill when applying a systems-based approach.

In nursing, integration of systems thinking is considered a key component of effective leadership and quality and safety practice¹¹⁻¹³ and a contemporary competency in nursing education as highlighted in the Future of Nursing report.³ Despite the systems thinking competency requirement in both medicine and nursing, an instrument to assess systems thinking for healthcare professionals is lacking.

A systematic review conducted in systems thinking and complexity science in health revealed that most published studies were conceptual with a significant lack of empirical and knowledge application findings.¹⁴ The ability to

J Gen Intern Med 35(8):2314-20

DOI: 10.1007/s11606-020-05830-1

© Society of General Internal Medicine 2020

Received September 8, 2019

Accepted March 30, 2020

Published online April 27, 2020

adequately measure systems thinking is needed and can be useful in education, practice, and research. For example, a measure of systems thinking may be vital to (1) increase our understanding of one of the mechanisms by which quality improvement processes achieve their results, (2) assist front-line staff to identify system contributions in quality improvement, (3) provide valuable information on the effectiveness of interventions, and (4) provide insight into the available human social capital on teams to help us understand why some teams are more effective than others. Additionally, given the premise that systems thinking can be taught and learned, the ability to assess systems thinking provides a way to measure the effectiveness of educational efforts.

Purpose

The purpose of this study was to develop, conduct a psychometric analysis, and determine the feasibility of a new measure of the extent to which individuals engage in systems thinking. We have named the instrument, the Systems Thinking Scale (STS).

Research Questions

- 1) What is the validity (content, construct, concurrent criterion-related, discriminate) of the Systems Thinking Scale?
- 2) What is the reliability (internal consistency and test-retest) of the Systems Thinking Scale?
- 3) What is the feasibility (time to complete and ease of use) of the Systems Thinking Scale?

Conceptual Framework

Two conceptual frameworks guided this project. First, Peter Senge's identification of systems thinking as one of five essential components of learning organizations provides a rationale for our focus on this concept.¹⁵ The linkages and feedback loops among the component parts maintain the interdependencies of the components of a system. Secondly, Batalden and Stoltz¹⁶ identify knowledge of a system as clarity about what, why, and how we make our products in healthcare. This includes understanding a healthcare system as a system of production that can be managed, and the ability to view, assess, and manipulate the interdependencies among the people, processes, products, and services. Both these frameworks recognize interdependencies as a key feature in systems thinking and provide the basis for the STS instrument development.

METHODS

Phase I: Definition and Item Generation

The definition and conceptual domains of systems thinking were identified using a grounded theory approach in a series of electronic email focus groups consisting of 11 international systems experts. In a group email format, experts provided

their responses to questions posed that included definitions and domains of systems thinking and dialogued on responses. Data from a series of iterative email discussions were analyzed for themes representing systems thinking domains and concept attributes. We also conducted an extensive literature search. Once the definition and domains were developed for each domain attribute of systems thinking (Fig. 1), experts reviewed the definition and potential scale items and validated if they were appropriate indicators of the domains (content validity). The experts were asked to identify other items not reflected in the item bank. Additionally, experts provided feedback on the stem of the question and response categories.

Phase II: Initial Field Testing of Items

Preliminary field testing for item clarity and feasibility of the instrument was conducted by administering the instrument to 10 interdisciplinary healthcare professionals. A project team member sat next to the individual while they completed the instrument and recorded any questions/comments about clarity of items, grammar, syntax, organization, appropriateness, and logical flow.

Phase III: Initial Psychometric Testing of Items

The revised 30-item scale then underwent initial item evaluation. The instrument was given to a sample of 390 healthcare professionals (returned surveys = 342, RR = 88%) and 200 healthcare professions students (returned surveys = 102, RR = 51%). Table 1 displays the demographic of the participants. A variety of professions (medicine, nursing, pharmacy, therapists) and levels of healthcare professionals (students, clinicians, and managers) participated from different delivery settings including primary care, specialty, emergency rooms, and surgical and critical care units.

Phase IV: Final Psychometric Testing

The remaining 26 items were included in a series of psychometric testing. Construct validity was assessed using principal axis factor analysis with oblique rotation to identify domains represented in the measure and further refine the set of items that comprise the STS instrument. The principal axis factoring analysis with oblique rotation was conducted for differing numbers of extracted factors, with a 3-factor solution appearing most useful. Reliability analyses were performed on each of the 3 identified factors using a coefficient alpha. Reliability was further assessed using test-retest in which 36 participants were administered the STS at a 2-week interval.

Discriminant validity was tested by comparing scores of known groups. In this situation, we compared a group of students who received systems thinking education during a 2-h course in the medical school (low dose, $N=78$), and a group of students in a 15-week quality improvement course that emphasized systems thinking (high dose, $N=11$), and a general health professionals course (nursing pharmacology,

<u>Working Definition:</u>
Systems thinking: The ability to recognize, understand, and synthesize the interactions, and interdependencies in a set of components designed for a specific purpose. This includes the ability to recognize patterns and repetitions in the interactions and an understanding of how actions and components can reinforce or counteract each other. These relationships and patterns occur at different dimensions: temporal, spatial, social, technical or cultural. It is fundamental to undertaking specific methodology or strategies to explore and redesign a set of components comprising a whole.
<u>Domains Derived from Expert Panel</u>
Sequence of events Causal sequence Multiple causations possible Variation of different types (random/special) Feedback Interrelations of factors Patterns of relationships

Figure 1 Systems thinking working definitions and domains.

$N = 32$) that conveyed no education specific to systems thinking. For all three groups, the STS was administered in a pretest-posttest format and validity assessed by calculating the difference in change between the groups. Comparing for quality improvement proficiency using the commonly used Quality Improvement Knowledge Application Tool (QIKAT) helped assess for concurrent criterion-related validity.¹⁷ For our analysis, during the first week of an interdisciplinary quality improvement course, students ($N = 11$) were administered the STS and the QIKAT. As a second concurrent, criterion-related validity test, we administered the STS to first-year medical students ($N = 78$) who received 4 h of systems thinking education related to an error case during their block one studies. Tests of correlations of the scores of the STS and their Summative Evaluation Questionnaire were done.

RESULTS

All phases in the development and testing of the STS were approved by the Case Western Reserve University and University Hospitals Institutional Review Board.

Table 1 Demographic Characteristics of Healthcare Professionals and Students in Phase 3

	Healthcare professional ($N = 342$) $N(\%)$	Students ($N = 102$) $N(\%)$
Administrator	10(3)	1(0.4)
Pharmacist	27(8)	0
Physician	55(16)	184(82.1)
Nurses	250(73)	27(12.2)
Public health	0	11(4.9)
Student	0	102(100)
	$\bar{X}(\text{SD})$	
Years working in healthcare	15.6(11.3)	
Years of QI work	6.6(8.5)	

Phase I: Definition and Item Generation

The experts identified 30 items in the initial phase, which is represented in Figure 2. Additionally the experts agreed upon the stem for the instrument: “When I want to make an improvement.” Response categories were “never” (0), “seldom” (1), “some of the time” (2), “often” (3), and “most of the time” (4). Figure 2 contains the initial item bank.

Phase II: Initial Field Testing of Items

In the initial field testing, it was found that the time to complete the instrument ranged from 5 to 10 min and all participants stated that it was easy to complete the survey. Revisions were made to the items and the instrument was formatted into a provisional form for psychometric testing. Changes made were in font size, minor rewording, and the replacement of “believe” to “think” in the instrument.

Phase III: Initial Psychometric Testing of Items

The initial psychometric testing of items revealed that mean item scores on the STS ranged from 0.76 (for item, “I consider the cause and effect that is occurring in a situation”) to 3.46 (for item “I include people in my work unit to find a solution”) and several items had means greater than 3.0. Four items (6, 22, 27, and 30) were excluded from all further analyses based on poor item-total correlations or poor item wording. The remaining 26 items were included in a series of psychometric testing.

Phase IV: Final Psychometric Testing

In evaluating the final 26-item STS, the results for construct validity using principal axis factoring analysis with oblique

1. I think the harder people work, the better the outcomes will be.
2. I seek everyone's view of the situation.
3. I look beyond a specific event to determine the cause of the problem.
4. I think understanding how the chain of events occurs is crucial.
5. I include people in my work unit to find a solution.
6. I think outcomes are random.
7. I think that lasting change relies on personal effort and motivation.
8. I think recurring patterns are more important than any one specific event.
9. I think of the problem at hand as a series of connected issues.
10. I consider the cause and effect that is occurring in a situation.
11. I consider the relationships among co-workers in the work unit.
12. I think that systems are constantly changing.
13. I propose solutions that affect the work environment, not specific individuals.
14. I focus on my first idea because it is often the best.
15. I keep in mind that proposed changes can affect the whole system.
16. The main reason for success is to get the person in charge to change.
17. I think my first impressions turn out to be very useful.
18. I think more than one or two people are needed to have success.
19. I keep the mission and purpose of the organization in mind.
20. I think small changes can produce important results.
21. I consider how multiple changes affect each other.
22. I focus primarily on the opinions of a champion in the system.
23. I think about how different employees might be affected by the improvement.
24. I think personal commitment is important in creating lasting change.
25. I try strategies that do not rely on people's memory.
26. I recognize system problems are influenced by past events.
27. I think that system-wide change is easy to accomplish.
28. I consider the past history and culture of the work unit.
29. I consider that the same action can have different effects over time, depending on the state of the system.
30. I think uncertainty and surprise are involved.

Figure 2 Systems Thinking Scale—initial item bank.

rotation indicated a 3-factor solution as most useful. Twenty items loaded high and cleanly on factor 1 (ranging from 0.35 to 0.69) which is interpreted as “interdependencies of system components.” Two items (1 and 7) loaded cleanly on a separate factor (0.67 and 0.44) interpreted as “personal effort.” Four items (14, 16, 17, 22) loaded on a distinct third factor, “reliance on authority,” although the loadings were not as high as for other items (0.53, 0.40, 0.44, 0.30). Factors were independent in that they did not correlate strongly with each other. Factors 1 and 2, however, correlated negatively (-0.21). The coefficient alpha for the 20-item subscale corresponding to factor 1 was 0.89. Coefficient alpha for the 2-item subscale corresponding to factor 2 was 0.55; this was reasonable given that it is a 2-item scale. Coefficient alpha for the 4-item subscale corresponding to factor 3 was 0.54, which was considered low for a 4-item scale.

Based on these reliability analyses and factor analyses, the 20 items of factor 1 (system interdependencies) appeared to be the strongest performing items of the intended construct of systems thinking (Fig. 3). Therefore, subsequent analyses were performed using the 20-item STS. Test-retest reliability revealed a correlation between the pre- and posttest of 0.74 (mean pretest 61.4, SD 6.5, and posttest scores, mean 59.1, SD 5.7). Discriminant validity of the 20-item STS was supported as the posttest significantly increased in the high-dose group only, and there was a significant difference in the STS scores between the high- and no-dose education groups (Table 2). Concurrent criterion-related validity was confirmed as scores on the QIKAT correlated with scores on the STS (Pearson correlation = 0.46). The correlation of the 20-item STS and the Summative Evaluation Questionnaire was low (Pearson Correlation = 0.28).

20-Item Systems Thinking Scale						
Instructions:						
Please read each of the statements and place an "X" in the answer box that indicates frequency of agreement with the statement:						
When I want to make an improvement. . .	Never	Seldom	Some of the time	Often	Most of the time	
1. I seek everyone's view of the situation.						
2. I look beyond a specific event to determine the cause of the problem.						
3. I think understanding how the chain of events occurs is crucial.						
4. I include people in my work unit to find a solution.						
5. I think recurring patterns are more important than any one specific event.						
6. I think of the problem at hand as a series of connected issues.						
7. I consider the cause and effect that is occurring in a situation.						
8. I consider the relationships among co-workers in the work unit.						
9. I think that systems are constantly changing.						
10. I propose solutions that affect the work environment, not specific individuals.						
When I want to make an improvement. . .	Never	Seldom	Some of the time	Often	Most of the time	
11. I keep in mind that proposed changes can affect the whole system.						
12. I think more than one or two people are needed to have success.						
13. I keep the mission and purpose of the organization in mind.						
14. I think small changes can produce important results.						
15. I consider how multiple changes affect each other.						
16. I think about how different employees might be affected by the improvement.						
17. I try strategies that do not rely on people's memory.						
18. I recognize system problems are influenced by past events.						
19. I consider the past history and culture of the work unit.						
20. I consider that the same action can have different effects over time, depending on the state of the system.						

Figure 3 20-item Systems Thinking Scale.

DISCUSSION

The STS is a 20-item user friendly, valid and reliable instrument that measures the systems thinking construct of system interdependencies in the context of quality improvement. As

the instrument takes 5–10 min to complete, it is a viable option for assessing systems thinking in both the clinical and education setting. Moreover, the instrument does not have reverse-coded items and thus allows one to easily calculate a final

Table 2 ANOVA Results for Differences among High, Low, and No Dose of Systems Thinking Education

Dose		Mean difference	Std. error	Sig. ^a	95% confidence interval for difference	
					Lower bound	Upper Bound
Low dose	High dose	-2.865	1.486	.05	-5.814	.084
	No dose	2.077	1.390	.14	-0.681	4.835
High dose	Low dose	2.865	1.486	.07	-0.084	5.814
	No dose	4.942*	1.856	.01	1.259	8.624

Based on estimated marginal means

^aAdjustment for multiple comparisons: least significant difference (equivalent to no adjustments)

*The mean difference is significant at the 0.05 level

score which decreases both respondent and evaluator burden and adds to its user friendliness.

One significant finding of the instrument development was how factor analyses independently revealed systems interdependencies as the major emerging construct. This corroborates previously published work suggesting interdependencies as the most prominent dimension of systems thinking. Specifically, Batalden and Mohr defined systems thinking as the discipline of seeing all elements in a given environment as interrelating.¹⁸ It considers the impact of actions in one part of the system on other parts of the system and views change as a process rather than a snapshot in time. The knowledge that the system is larger than the sum of its parts and recognizing connections within the system is critical to being a systems thinker, which is what is being assessed in the 20-item STS.

Although our psychometric evaluation supported system interdependencies, it did not support factors 2 (personal effort) and 3 (reliance on authority). Although these dimensions may be important, the items, as worded and based on item responses from this particular sample, did not function well as part of a single overall measure of systems thinking. This could be a function of the quality improvement focus of STS and these factors not being highlighted in that context. Further studies are needed to continue to develop these components of systems thinking, as they were identified by the international expert panel in phase 1 of our study.

A strength of the STS is that the final 20-item instrument included sufficient representation of the initial item bank as generated by the international expert team who had a wide range of expertise and experience in systems thinking and quality improvement. Another strength is that during all phases of the psychometric testing, we ensured that we administered the instrument to a wide variety of interprofessional team members and learners who came from different clinical settings and QI experience. This is critical as it speaks to the generalizable use of the STS as it can be used in any setting with different healthcare professionals regardless of discipline, level of learning, or expertise.

Our study has limitations. First, the 20-item STS is based on self-report and does not measure actual observed competence. The self-report nature of an individual’s perspective limits its application to any extrapolation of a “team” measure of systems thinking. One potential way to get around this is to use the STS in a 360-degree fashion with other team members and see if there is consistency in findings. Future research to derive a “mean score” for the team using an average calculation across all members and testing this score against a criterion of successful QI would be useful. We will need to test several approaches using different scoring options for obtaining a team score and determine which approach appears most suitable and performs best in our series of psychometric testing. An additional limitation is that we were unable to establish concurrent validity with the STS. We believe that currently there is not an appropriate “gold standard” measure of systems thinking with which to compare our new instrument. For our psychometric evaluation of concurrent validity, we used the QIKAT and the Structured Evaluation Questionnaire and found low correlations. The low correlations are likely due to the poor match of the comparison concepts as systems thinking and interdependencies are not key constructs of the QIKAT and Structured Evaluation Questionnaire. Future research to evaluate concurrent validity with new systems thinking instruments from other disciplines is warranted.¹⁹

Systems thinking is a critical component of quality improvement learning and the foundation for systems-based practice (SBP), a core competency mandated by the ACGME for physicians in training. SBP has been a challenging competency to assess in medicine for a variety of reason. Specifically, there are no prior reliable tools for the foundation of this competency, systems thinking. We believe the 20-item STS could be a feasible method of assessment that not only becomes part of a learner portfolio but also helps to track progress over time especially in longitudinal clinical settings. Systems-based practice also includes leadership and operations competencies; it is possible that the STS may be adapted to capture these additional components.

Furthermore, recent literature on health systems science as the third pillar of medical education underscores the need for systems thinking which is considered the “interlinking domain” of Health Systems Science and provides the foundation for Health Systems Science competencies.¹⁰

Future research is needed to continue to assess the psychometrics of the current 20-item STS. Discriminant validity evidence would be enhanced by testing additional sets of learners at other institutions and/or at different levels of training and capturing qualitative data using structured interviews. Predictive validity evidence would be enhanced by testing if systems thinking might improve with practical experience in QI operations and perhaps mindful clinical practice even in the absence of formal training. The use of the STS in improvement science research might include the relationship of team members’ scores on the STS and quality improvement success, the effect of increasing systems thinking and quality improvement

staff engagement, and the relationship between a microsystem improvement culture and degree of systems thinking. Future educational research is needed to further evaluate educational strategies to improve systems thinking.

Since the development of the STS, there are few studies that used the Scale. Nurses with higher levels of systems thinking reported higher safety culture and less medication errors,²⁰ were more likely to report medical errors and less likely to experience the occurrence of adverse events through a self-report question,²¹ and reported higher levels of safety attitude, knowledge, and skill.^{22, 23} Improvement in systems thinking is reported in both nurses²⁰ and medical students.²⁴

In summary, the 20-item STS is a valid and reliable instrument that is easy to administer and takes less than 10 min to complete. Further research using the STS has the potential to advance the science and education of quality improvement in two main ways: (1) increase understanding of a critical mechanism by which quality improvement processes achieve results, and (2) evaluate the effectiveness of our educational strategies to improve systems thinking. Demonstrating skill in systems thinking is a goal for healthcare professionals to effectively engage in successful quality improvement.²⁵

Corresponding Author: Mary A. Dolansky, PhD, RN; Frances Payne Bolton School of Nursing Case Western Reserve University, 10900 Euclid Ave., Cleveland, OH 44106, USA (e-mail: Mary.dolansky@case.edu).

Funding Information Funding for the study was provided by the Robert Wood Johnson Foundation improving the Science of Continuous Quality Improvement Program and Evaluation.

Compliance with Ethical Standards:

Conflict of Interest: The authors declare that they do not have a conflict of interest.

REFERENCES

1. Accreditation Council for Graduate Medical Education (2007). website [On-line].
2. Institute of Medicine. Committee on the Robert Wood Johnson Foundation Initiative on the Future of Nursing. The Future of Nursing: Leading Change, Advancing Health. Washington, DC: National Academies Press; 2011.
3. Dolansky, M. A., & Moore, S. M. (2013). Quality and safety education for nurses (QSEN): The key is systems thoughtful. *Online Journal of Issues in Nursing*, 18(3).
4. Calhoun, JG, Ramiah, K, Weist, EM, & Shortell, SM. Development of a core competency model for the master of public health degree. *American Journal of Public Health*. 2008; 98: 1598-1607.
5. Frenk, J, Chen, L, Bhutta, ZA, Cohen, J, Crisp, N, Evans, T, & Kistnasamy, B. Health professionals for a new century: transforming education to strengthen health systems in an interdependent world. *Lancet*. 2008; 376:1923-1958.
6. Vincent, C, Taylor-Adams, S, & Stanhope, N. Framework for analyzing risk and safety in clinical medicine. *BMJ*: 199;316:1154.
7. Carayon, P, Hundt, AS, Karsh, BT, Gurses, AP, Alvarado, CJ, Smith, M, & Brennan, PF. Work system design for patient safety: the SEIPS model. *BMJ Quality & Safety* 2008; 15:i50-i58.
8. Bowie, P, McNaughton, E, Bruce, C. Enhancing the effectiveness of significant event analysis: exploring personal impact and applying systems thinking in primary care. *The Journal of continuing education in the health professions*. 2016; 36:195.
9. Phillips, JM, Stalter, AM, Winegardner, S, Wiggs, C, & Jauch, A. Systems thinking and incivility in nursing practice: An integrative review. *Nursing forum*. 2018; 5: 286-293.
10. Gonzalo, JD, Dekhtyar, M, Starr, S; et al. Health Systems Science Curricula in Undergraduate Medical Education: Identifying and Defining a Potential Curricular Framework. *Academic Medicine*. 2017;92:123-131.
11. Bleich, MR. Developing leaders as systems thinkers—Part I. *The Journal of Continuing Education in Nursing*. 2014; 45:158-159.
12. Phillips, JM, Stalter, AM, Dolansky, MA, & Lopez, GM. Fostering future leadership in quality and safety in health care through systems thinking. *Journal of Professional Nursing*. 2016; 32:15-24.
13. Phillips, J M, & Stalter, AM. Integrating systems thinking into nursing education. *The Journal of Continuing Education in Nursing*. 2016; 46:395-397.
14. Rusoja, E, Haynie, D, Sievers, J, et al. Thinking about complexity in health: A systematic review of the key systems thinking and complexity ideas in health. *Journal of evaluation in clinical practice*. 2018; 24:600-606.
15. Senge, P. *Discipline: The art and practice of the learning organization*. New York: Doubleday/Currency; 1990.
16. Batalden, PB & Stoltz, PK. A framework for the continual improvement of health care: building and applying professional and improvement knowledge to test changes in daily work. *Jt.Comm J.Qual.Improv*.1993; 19:424-447.
17. Morrison, L, Headrick, L, Orgrinc, G, & Foster, T. The quality improvement knowledge application tool: An instrument to assess knowledge application in practice-based learning and improvement. *Journal of General Internal Medicine*. 2003;18:250.
18. Batalden, PB, & Mohr, JJ. Building knowledge of health care as a system. *Quality Management in Health Care*. 1997; 5:1-12.
19. Randle, J. M., & Stroink, M. L. (2012). Systems thinking cognitive paradigm and its relationship to academic attitudes and achievement. *Unpublished honors thesis*. Lakehead University, Thunder Bay.
20. Tetuan, T., Ohm, R., Kinzie, L., McMaster, S., Moffitt, B., & Mosier, M. (2017). Does Systems Thinking Improve the Perception of Safety Culture and Patient Safety? *Journal of Nursing Regulation*, 8(2), 31-39.
21. Hwang, J. I., & Park, H. A. (2017). Nurses' systems thinking competency, medical error reporting, and the occurrence of adverse events: A cross-sectional study. *Contemporary nurse*, 1-11.
22. Dillon-Bleich, K. (2018). *Keeping Patients Safe: The Relationships Among Structural Empowerment, Systems Thinking, Level of Education, Certification and Safety Competency* (Doctoral dissertation, Case Western Reserve University).
23. Masoon, A. (2019). Relationships among systems thinking, safety culture, safety competency, and safety performance of registered nurses in Saudi Arabia. Doctoral dissertation, Case Western Reserve University).
24. Aboumatar, H. J., Thompson, D., Wu, A., Dawson, P., Colbert, J., Marsteller, J., ... & Pronovost, P. (2012). Development and evaluation of a 3-day patient safety curriculum to advance knowledge, self-efficacy and system thinking among medical students. *BMJ Qual Saf*, 21(5), 416-422
25. Plack, MM, Goldman, EF, Scott, et al. Systems thinking and systems-based practice across the health professions: An inquiry into definitions, teaching practices, and assessment. *Teaching and Learning in Medicine*. 2018;30:242-254.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.