

**Testing Nursing Students' Clinical Judgment in a Patient Deterioration Simulation**  
**Scenario: Development of a Situation Awareness Instrument**

Patrick Lavoie, RN, Ph.D. candidate, Faculty of Nursing, Université de Montréal, Montréal, Canada

Sylvie Cossette, RN, Ph.D., Vice-dean, research and international relations, Faculty of Nursing, Université de Montréal, Montréal, Canada.  
Researcher at the Montreal Heart Institute, Montréal, Canada.

Jacinthe Pepin, RN, Ph.D., Professor and Director of the Center for Innovation in Nursing Education (CIFI), Faculty of Nursing, Université de Montréal, Montréal, Canada.

The first author received doctoral scholarships from the Quebec Nursing Intervention Research Network (RRISIQ), funded by the *Fonds de recherche du Québec – Santé*, from the *Fonds de recherche du Québec – Société et Culture (FRQ-SC)*, and from the *Ministère de l'Éducation, de l'Enseignement supérieur et de la Recherche du Québec*. The *Équipe FUTUR*, funded by the *FRQ-SC*, funded the editing of this paper.

The authors would like to acknowledge the valuable contribution of Dr. Sébastien Béland from the Faculty of education, Université de Montréal, for his valuable contribution to the planning of the statistical analyses for this study.

Corresponding author: Patrick Lavoie, [patrick.lavoie.1@umontreal.ca](mailto:patrick.lavoie.1@umontreal.ca)

Coauthors: [sylvie.cossette.inf@umontreal.ca](mailto:sylvie.cossette.inf@umontreal.ca), [jacinthe.pepin@umontreal.ca](mailto:jacinthe.pepin@umontreal.ca)

Work address for all authors: Faculté des sciences infirmières, Université de Montréal, C.P. 6128, succ. Centre-Ville, Montréal, QC, H3C 3J7

## ABSTRACT

**Background:** Situation awareness may be used to operationalize nursing students' clinical judgment of patient deterioration simulation scenarios.

**Objectives:** To develop and test an instrument to measure bachelor-level nursing students' situation awareness in a patient deterioration simulation scenario, using the Situation Awareness Global Assessment Technique (SAGAT).

**Design:** Instrument development and validation.

**Settings:** A faculty of nursing of a French-Canadian university.

**Participants:** 15 critical care experts and 234 bachelor-level nursing students from a critical care course.

**Methods:** The queries were developed from evidence and guidelines regarding nurses' assessment and response to patient deterioration and an inventory of nursing diagnosis. After expert content validation, the instrument was administered to three cohorts of nursing students in a high-fidelity simulation with a scenario of hypovolemic hemorrhagic shock. Difficulty, discrimination, and fidelity indices were computed. The impact of the instrument on student's performance was assessed with a post-simulation survey.

**Results:** The instrument comprised 31 queries, which obtained high content validity indices. Most showed satisfying difficulty, discrimination, and fidelity properties. Inadequate properties of the queries may be explained by the content of the simulation scenario, the assessment practices of nursing students, and their reliance on medical assistance. Students perceived that completing the instrument helped them realize what they forgot to assess in the simulation.

**Conclusions:** This instrument appears as a promising research tool, although it still needs to be tested with other populations and in other patient deterioration simulation scenarios.

**Keywords:** Patient deterioration; research instrument; instrument validation; test construction; reliability and validity; simulations; students, nursing; situation awareness

## INTRODUCTION

Nurses need to notice and interpret signs of deterioration in hospitalized patients to reach sound clinical judgment (Department of Health, 2009; Liaw et al., 2011; Odell et al., 2009). Clinical judgment consists of a nurse's understanding of a patient's situation and the decisions on actions to be taken (Tanner, 2006). As the outcome of the decision-making process, clinical judgment accuracy depends on nurses' assessment and analysis of a patient's situation.

Educators have used simulations to prepare nursing students to recognize and respond to patient deterioration situations (Fisher and King, 2013). However, they mostly assessed learning outcomes other than clinical judgment using scales of participants' satisfaction or self-rating of their confidence, skills, or competence. Although it is important to consider participants' reactions to teaching, these measures do not provide evidence of its effectiveness for clinical judgment. Others measured participants' knowledge or their skills in simulations with objective structured clinical examination checklists. These approaches are questionable when it comes to cognitive processes where knowledge is necessary but at the same time insufficient to ensure proper understanding and decision making (Tanner, 2011). Besides, observation of actions during a simulation may provide data on the decisions the nurses made but not on their understanding of a patient's situation.

Situation awareness, a concept from military aviation exported to human factors engineering, is defined as "knowing what is going on around you" (Endsley, 2000b). This theoretical construct consists of three levels: perception of cues, comprehension of their meaning, and projection of their status to the near future (Endsley, 2000b). When applied to nursing, it was defined as a nurse's perception of relevant clinical cues, comprehension of their meaning, and projection of required interventions (Sitterding et al., 2012). By extension, situation awareness is a concept that could describe the process that surrounds clinical judgment: noticing, interpreting, and responding to a patient's needs or status (Tanner, 2006). Therefore, a measure of nurses' situation awareness may be a way to operationalize clinical judgment in a simulation. This paper reports on the development of an instrument to measure bachelor-level nursing students' situation awareness in a patient deterioration simulation scenario and its testing for reliability and validity.

### **Measurement of situation awareness: SAGAT**

The Situation Awareness Global Assessment Technique (SAGAT) allows for direct measurement of situation awareness by assessing an individual's perception of a situation instead of inferring it from the observation of the individual's behaviors. This technique was first

developed in the context of aircraft engineering to provide an objective and valid measure of all three levels of situation awareness that would not alter situation awareness in the process of measurement, as other available methods were likely to do (Endsley, 1995). It consists of a series of short queries about an individual's perception, comprehension, and projection of a simulated situation at a precise moment (Endsley, 2000a). The content of these queries is determined after an analysis of the information required to complete a task during the simulation. The individual is queried during a brief freeze in the simulation while all data from the simulator are hidden from the individual.

To use the SAGAT in a patient deterioration simulation scenario, we needed queries designed for scenarios where a patient was exhibiting signs and symptoms of hemodynamic instability attributable to a worsening clinical state. Through database search, we retrieved two lists of queries that had been used with medical students in trauma and sepsis simulation scenarios (Hansel et al., 2012; Hogan et al., 2006). We retrieved another list of queries (McKenna et al., 2014) that was used in studies with nurses or nursing students experiencing scenarios such as acute myocardial infarction, hypovolemic shock, septic shock, or chronic obstructive pulmonary disease (Cooper et al., 2013; Cooper et al., 2010; Cooper et al., 2011; Endacott et al., 2010). Only Hogan et al.'s (2006) list was tested for fidelity and validity with a sample of 16 participants. With this small sample, it yielded a Cronbach's alpha of 0.77 and a Pearson's correlation of 0.81 when compared to a score on a checklist of assessment and management steps required in the scenarios where it was tested. Overall, these three lists of queries only addressed some of the early warning signs of patient deterioration described in the literature. Some open-ended queries yielded only qualitative data that would need to be quantified in order to obtain a situation awareness score. Therefore, we used them as inspirations for the development of a comprehensive list of SAGAT queries to be used in a patient deterioration simulation scenario.

## MATERIAL AND METHODS

This was an instrument development and validation study embedded in a mixed methods research project. It aimed at developing an instrument to be used with the SAGAT in order to measure nursing students' situation awareness in a patient deterioration simulation scenario and to test its psychometric properties in a sample of bachelor-level students in a critical care course.

### **Context**

Our simulations took place in individual rooms designed to replicate the hospital environment and involved high-fidelity (male patient) manikins. Real medical equipment and supplies (e.g., cardiac monitor, oxygen supply, medication) were available to maximize realism. The scenarios followed three phases: initial assessment, deterioration, and stabilization of the simulated patient. An operator controlled the evolution of the scenarios depending on students' completion of expected tasks. Students participated in groups of five to six.

Since we aimed at measuring nurses' situation awareness *prior* to the patient's deterioration in the second phase, we decided that the simulation would be frozen at the end of the first phase of the scenario, that is, after students completed an initial assessment. During the freeze, the operator would hide the bedside monitor and stop the manikin's breathing. Then, another person who administered the questionnaire would handout a pen-and-paper version of the instrument and read each query as a mean to pace the completion of the questionnaire. Once it was completed and picked up, the operator would reveal the hidden data and resume the simulation with the freeze planned to last for 180 seconds.

### **Development of the queries**

To develop the queries, we analyzed the information nurses need in patient deterioration situations. For the first theoretical level of situation awareness (perception), we relied on clinical guidelines and standards for assessment, recognition, and response to acute illness (National Institute for Health and Clinical Excellence, 2007; Royal College of Physicians, 2012) to identify objective quantifiable physiological signs of deterioration. We also used results of qualitative studies on how nurses recognize deteriorating patients through subtler and less quantifiable changes in their condition (Cioffi et al., 2009; Gazarian et al., 2010). The signs that could not be directly simulated on our high-fidelity manikin (e.g., sweat or arterial blood gases) were not considered. For the second level (comprehension), we examined an inventory of nursing diagnoses (NANDA International, 2012) to identify those which nurses could arrive at through a critical care assessment approach: the Airway-Breathing-Circulation-Disability-Exposure ABCDE survey (Emergency Nurses Association, 2007). For the third level (projection), we kept queries that were used in studies with nurses or nursing students experiencing patient deterioration simulation scenarios (McKenna et al., 2014). To those, we added one query that was based on nurses frequent reporting of a "gut feeling" that something bad was about to happen to the patient and another about communicating with medical staff to obtain assistance (Odell et al., 2009).

The first list consisted of 36 queries, with most ( $n = 22$ ) requiring Yes/No answers. At the perception level, five queries asked about the simulated patient's vital signs. For those, we followed Hogan's (2006) proposal by permitting a range of  $\pm 10\%$  around the correct answer. For instance, for the heart rate query, a variation of  $\pm 12$  around a value of 120 bpm would be considered correct. The first-level query on level of consciousness required an answer on the Alert-Verbal-Painful-Unresponsive (AVPU) scale. For the projection level, queries that addressed vital signs were multiple-choice questions with three possible answers: "will go up," "will go down," or "will stay the same." An additional possible answer, "I don't know" was added to all queries assessing the three levels.

### **Content validation**

*Sample.* We followed recommendations for content validation considering adjustment for chance agreement (Polit et al., 2007). A large panel of experts ( $n = 8-12$ ) was needed for the first round and a smaller panel ( $n = 3-5$ ) participated in the second one. We recruited 15 critical care experts through a network sampling procedure (Grove et al., 2013) among Quebec's universities and their associated clinical institutions. Their socio-demographic data are presented in Table 1.

*Procedure.* Through an emailed questionnaire, the 11 experts of the first validation round were asked to rate the relevance of the queries on 4-point Likert scales, from not relevant to highly relevant to their respective theoretical level of situation awareness (Polit et al., 2007). They also acknowledged if the queries were clearly formulated, exclusive to their theoretical level, and if they were necessary or should be omitted. Space was available for comments or suggestions of concepts that were absent. The second content validation round followed the same procedure with eight experts. In both rounds, we asked the experts to rate the acceptability of the range allowing correct numerical answers.

*Data analysis.* Three indices were computed for each query. The item content validity index (I-CVI) represented the proportion of experts who rated the query as relevant or highly relevant to its theoretical level. According to Polit et al. (2007), an acceptable I-CVI should exceed 0.78 when eight or more experts are consulted. We followed the same rationale and criteria to compute the index of item clarity (I-CI) and the index of necessity (I-NI). After the second round, we computed the scale content validity index by averaging I-CVI across items (S-CVI/Ave), which should reach 0.90 (Polit et al., 2007).

In both rounds, the criteria for exclusion of a query were (a) an I-CVI  $< 0.78$ , (b) an I-NI  $< 0.78$ ,

or (c) non-exclusivity reported by at least two experts with relevant explanations. We examined all experts' suggestions for reformulations or absent concepts that should be included in the instrument and adjusted the queries accordingly. For the acceptability of the range for correct numerical answers, we aimed at consensus among experts.

### **Difficulty, discrimination, reliability, and perceived impact on performance**

*Sample.* The version of the instrument that was achieved after content validation was administered to three cohorts of baccalaureate nursing students in a scenario where a patient experienced hypovolemic shock secondary to an arterial bleeding from a femoral coronarography site. A convenience sample was formed with students from a critical care course delivered on both campuses of our university. Every student following the course during autumn 2014 (cohort 1 from campus 1) and winter 2015 (cohort 2 from campus 1 and cohort 3 from campus 2) was invited to participate in the study. The university's ethical review board approved this study and informed consent was obtained from all participants.

*Procedure.* Students participated in the scenario during two days on the sixth, fifth, or eighth week of their course depending on their cohort. The administration of the questionnaire followed the procedure detailed in the Context section of this paper. At the time of the scenario freeze, the operator also completed the questionnaire to provide the scoring key based on the actual values reported on the bedside monitor. After the simulation, participants were asked to complete a socio-demographic questionnaire.

To appraise impact of the instrument on their performance, students from the first cohort completed a short survey consisting of 4-point Likert scales to rate the magnitude of its influence and its impediment to their performance. Sections were available for participants to comment.

*Data analysis.* The instrument is composed of 31 queries divided in three levels of situation awareness: (1) perception, (2) comprehension, and (3) projection. Queries from the first level were subdivided in two subsets: (1a) queries on the perception of objective signs of deterioration and (1b) queries on the perception of less quantifiable signs of deterioration. For each query, we computed the difficulty index ( $p$ ), discrimination index ( $D$ ), and reliability index ( $\rho_{pbis-i}$ ) (Crocker and Algina, 2008; Laveault and Grégoire, 2014). The difficulty index ranges from 0 to 1 and represents the proportion of students who answered a query correctly. A high  $p$ -value indicates that a query is easy to answer.

To compute the discrimination and reliability indices, a criterion of reference is needed. We

used students' scores on the level (or subset) of the instrument for each query included therein. For instance, query 3 (respiratory rate) is a query from level 1, subset 1a. Accordingly, the discrimination and reliability indices of this query were calculated using individual score (for the discrimination index) or mean score (for the reliability index) for subset 1a.

To compute the discrimination index of a query, we formed two contrasted groups based on students' individual scores on the level of that query (higher 30% and lower 30%). The  $D$ -value was calculated by subtracting the  $p$ -value of the query in the lower group to its  $p$ -value in the higher group ( $p_{upper} - p_{lower}$ ). The  $D$ -value ranges from 0 to 1 and increases as a query better discriminates between the two groups. As for the reliability index, it represents the correlation between the mean score on a query and the mean score on its respective level or subset with the query tested removed from the level's score.

The indices were computed with IBM® SPSS® Statistics Version 21 and Microsoft® Excel® for Mac 2011. We used arbitrary but plausible thresholds to screen for difficult ( $p < 0.30$ ) and easy items ( $p > 0.80$ ). We considered Ebel's (1965) proposed criteria for item revision ( $0.20 < D < 0.40$ ) and elimination ( $D < 0.20$ ). Based on the sample size, the minimum acceptable value for  $\rho_{pbis-i}$  was set at 0.13, which equals to two standard errors above zero (Crocker and Algina, 2008). Furthermore, we tentatively computed the Kuder-Richardson Formula 20 (KR-20), an indicator similar to Cronbach's alpha for dichotomous answers, even if it is likely to be an underestimation of its score reliability since the actual form of the instrument does not respect the assumptions of the tau-equivalent model of measurement (Graham, 2006). The survey addressing students' perception of the impact of the instrument on their performance was used for descriptive purposes.

## RESULTS

### **Content validation**

Following the first round of expert content validation, six queries were kept as is, 19 were reformulated, and ten were removed. One level-one query was added (pulse regularity) along with two level-two queries (infection and delirium) and five level-three queries (projection of respiratory rate and diuresis, informing the doctor, doctor at the bedside, and initiating an emergency intervention). The second version of the instrument comprised 33 queries (14 first-level, 11 second-level, and 8 third-level). Experts asked that the range for correct numerical answers be reduced and not expressed as percentage but in the unit of the physiological parameter. For instance, instead of



allowing a variation of  $\pm 10\%$  from the actual heart rate, we permitted a range of  $\pm 5$  bpm.

In the second round, the query about projection of diuresis was removed, since experts considered that a change in diuresis would only occur after a period longer than the time frame of the simulation. Although it did not reach a sufficient I-CVI, one query (asking for the physician's presence at the bedside) was kept for exploratory purpose. The delirium query had perfect I-CVI but was excluded, since qualitative comments suggested that it was rather a neuropsychiatric than hemodynamic issue. Most experts deemed the emergency intervention query too vague. We made it more precise by specifying the administration of a fluid bolus as an emergency intervention. All experts judged the new acceptability range to be adequate. The "gut feeling" query was removed in the first round because of insufficient I-CVI.

The full instrument is presented in Table 2. All queries reached I-CVI, I-CI, and I-NI between 0.86 and 1.00, except for exploratory queries. Considering that one expert refrained from evaluating four queries, this means that at least six out of seven or seven out of eight experts rated them as relevant, clearly formulated, and necessary. This version of the scale reached an S-CVI/Ave of 0.97 without the two queries that were not rated by experts.

This study was conducted with the French version of the scale. The English version was obtained through a translation and back-translation process (Streiner and Norman, 2008). A bilingual clinical nurse specialist translated the original instrument. An independent certified translator translated it back to French with few adjustments needed to correspond to the original version.

### **Difficulty, discrimination, reliability, and perceived impact on performance**

*Sample.* Of the 294, 136, and 143 students of the three cohorts, we collected data from 109, 77, and 48 students. Their socio-demographic data are presented in Table 3.

*Results.* Participants' answers were scored with one point for a correct answer and zero points for an incorrect or an "I don't know" answer. Query 1 (blood pressure) was divided into two queries (1.1 and 1.2) to isolate systolic from diastolic blood pressure. Students' total mean score reached 21.64 of 32 (SD = 3.94). Two-tailed independent samples T-tests showed no significant differences between total mean scores for all cohorts: cohort 1 (M = 21.28, SD = 3.64) and cohort 2 (M = 21.61, SD = 4.16);  $t(184) = -0.58$ ,  $p = 0.56$ ; cohort 1 and cohort 3 (M = 22.42, SD = 4.36);  $t(155) = -1.70$ ,  $p = 0.09$ ; cohort 2 and cohort 3;  $t(123) = -1.03$ ,  $p = 0.30$ . This showed that cohorts' scores were comparable. The mean for each level and the difficulty, discrimination and reliability indices of the queries for the pooled sample are presented in Table 4.

Twelve queries exceeded the threshold for easy queries. This applied to three first-level queries: query 5 (level of consciousness,  $p = 0.86$ ) of the objective signs, and queries 11 (agitation,  $p = 0.89$ ) and 12 (unusual pain,  $p = 0.84$ ) of the less quantifiable signs. At the second level, queries 15 (airway patency,  $p = 0.92$ ), 21 (neurological involvement,  $p = 0.90$ ), and 23 (infection,  $p = 0.84$ ) were considered easy. All third-level queries exceeded the easy threshold except for queries 24 (bolus administration,  $p = 0.55$ ) and 26 (doctor at bedside,  $p = 0.26$ ). For the whole scale, only query 3 (respiratory rate) was difficult ( $p = 0.24$ ).

Easy queries from the two first levels also showed poor discriminatory values with  $D$  smaller than 0.25 except for query 23 (infection,  $D = 0.44$ ). Second-level queries 18 (peripheral perfusion,  $D = 0.07$ ) and 22 (bleeding,  $D = 0.27$ ) had medium-range  $p$ -values but inadequate  $D$ -values. Only third-level query 25 (advise doctor,  $D = 0.18$ ) showed an insufficient  $D$ -value. Reliability indices of first-level queries 5, 12, and 13 did not reach a satisfactory level ( $\rho_{pbis-i} < 0.13$ ). The same applies to second-level queries 18 (peripheral perfusion) and 22 (bleeding) and third-level query 26 (doctor at bedside). For the whole scale, KR-20 reached 0.64.

Figure 1 presents the relationship between  $p$  and  $D$ , with horizontal lines marking Ebel's (1965) criteria for  $D$  and vertical lines representing thresholds for  $p$ . Most queries ( $n = 21$ , 65.6%) appear above the threshold for proper discrimination and have ideal medium-range  $p$ -values. However, queries 5, 11, 12, and 25 have high  $p$ -values, which influence their discriminatory indices below a minimum acceptable value ( $D < 0.20$ ). Likewise, queries 15, 21, 28, and 31 have high  $p$ -values, which lowers their  $D$ -value; however, in this case, a lower  $D$ -value ( $0.20 < D < 0.40$ ) is tolerable. Queries with medium-range  $p$ -values and unsatisfactory  $D$ -values (i.e., queries 18, 22, and 26) should be examined regarding other criteria, such as their reliability indices, which have shown to be insufficient in those three cases.

In the survey on the impact of the instrument on performance, students ( $n = 109$ ) reported that it influenced their performance tremendously ( $n = 8$ ), a lot ( $n = 24$ ), slightly ( $n = 44$ ), or not at all ( $n = 33$ ). Analysis of comments revealed that completing the questionnaire helped them realize what they forgot in their initial assessment of the situation and gave them a pause to reflect on what was going on with the patient. This was perceived as a positive influence on their actions in the latter phases of the scenario. However, even if they felt its influence, most reported that the instrument did not impede their performance ( $n = 93$ ). Some students commented that it lowered their anxiety level in the simulation.

## DISCUSSION

This paper presents the development and testing of a list of SAGAT queries to measure nursing students' awareness in a patient deterioration simulation scenario. All queries reached high indices of content validity. However, certain queries had indices below threshold for proper discrimination. Figure 1 explains how high  $p$ -values influence some discriminatory indices in a downward direction (e.g., queries 15 and 21). Since most students answered them correctly, these queries are unlikely to attain high discriminatory qualities. Queries with  $D$ -values below 0.20 should be examined cautiously since low-performing and high-performing students have a similar probability to answer them correctly. For instance, the information necessary to obtain correct answers for queries 5 (level of consciousness), 11 (agitation), and 12 (unusual pain) may be more evident in comparison to other first-level queries. Since the "patient" (manikin) was alert in the scenario and complained loudly about pain without being agitated, participants could determine the answers to these queries while completing the questionnaire, even if they did not intentionally assess these parameters while participating in the scenario. This might explain why these items differed from those where the "patient" needed to be assessed purposely and why they had the highest  $p$ -values of the first level.

Query 18 (peripheral perfusion) was the poorest discriminator of the instrument and had one of the lowest correlations to its level score along with query 22 (bleeding). This is intriguing, given the simulation scenario where the patient was experiencing an arterial bleeding that manifested through an alteration of his right leg's perfusion. It appears that students could answer queries on bleeding and peripheral perfusion no matter how they performed on their comprehension of other aspects of the simulation. This might be symptomatic of novice nurses' tendency to focus on one problematic aspect of a patient situation and to miss other relevant cues, which has been described in the literature (Benner et al., 1992; Hoffman et al., 2009; O'Neill et al., 2005).

Students obtained their highest subscore in the third level (projection, 77%), with queries 27 to 31 all having  $p$ -values higher than 0.80 and reaching proper discriminatory values. These results are comparable to those of other studies (Bogossian et al., 2014; Cooper et al., 2013; McKenna et al., 2014). However, given the patient's condition at the end of the first phase of the scenario, it was necessary to advise the doctor (query 25) but not to ask him to come to the bedside (query 26). Most students (91%) replied that they would call the doctor, and more than half of them (54%) asked him to come to the bedside. These results show how heavily students rely on medical

assistance, even if not needed, during a patient deterioration simulation scenario, possibly because of their inexperience with such situations. Otherwise, for all levels of situation awareness,  $\rho_{pbis-i}$  values showed that queries that were already identified as problematic with other indices demonstrated poor relationships to their theoretical level.

Individual measurement of situation awareness might be problematic when groups of students participate in the simulations. Because students divide tasks among their team members, individual assessment errors may affect others' test performance. For instance, when students measured respiratory rate, it was observed that they frequently obtained erroneous results that they shared with their colleagues. This was confirmed with query 3 (respiratory rate), which ended up being the most difficult. This also reflects existing evidence that respiratory rate is often neglected in nurses' assessment practices (Hogan, 2006; Ludikhuizen et al., 2012; Pantazopoulos et al., 2012). Besides, the level of participation of students in the simulation might have an impact on their score on the instrument. Since we did not quantify students' participation in the simulations, this hypothesis will need to be tested in a future study.

According to the psychometrical results, queries from level one (objective signs: 5; less quantifiable signs: 11, 12), level two (18, 22), and level three (25, 26) should be excluded or revised. However, we would recommend their inclusion since they obtained high indices of content validity and their removal only resulted in a negligible increase of KR-20 to 0.66. Our results might reflect participants' inexperience with patient deterioration situations. Further research is needed to study this hypothesis by administering the instrument to expert nurses in the same simulation scenario. There is also a strong possibility that the queries might react differently in another patient deterioration simulation scenario.

## CONCLUSION

We have developed a list of SAGAT queries, which could be used to assess nursing students' situation awareness in a patient deterioration simulation scenario, and tested their validity and reliability. Even if the list still needs to be revised and improved, it appears to be a promising tool for researchers who wish to assess students' clinical judgment in such scenarios. Educators who want to increase students' awareness of what they forgot in their initial assessment and what might happen to the patient could also use this instrument. Further development and testing will be needed prior to its use as a certification tool.

## REFERENCES

- Benner, P., Tanner, C., Chelsa, C., 1992. From beginner to expert: Gaining a differentiated clinical world in critical care nursing. *Advances in Nursing Science* 14 (3), 13-28.
- Bogossian, F., Cooper, S., Cant, R., Beauchamp, A., Porter, J., Kain, V., Bucknall, T., Phillips, N.M., Team, F.A.R., 2014. Undergraduate nursing students' performance in recognising and responding to sudden patient deterioration in high psychological fidelity simulated environments: An Australian multi-centre study. *Nurse Education Today* 34 (5), 691-696. DOI: 10.1016/j.nedt.2013.09.015
- Cioffi, J., Conway, R., Everist, L., Scott, J., Senior, J., 2009. 'Patients of concern' to nurses in acute care settings: A descriptive study. *Australian Critical Care* 22 (4), 178-186. DOI: 10.1016/j.aucc.2009.07.001
- Cooper, S., Cant, R., Porter, J., Missen, K., Sparkes, L., McConnell-Henry, T., Endacott, R., 2013. Managing patient deterioration: Assessing teamwork and individual performance. *Emergency Medicine Journal* 30, 377-381. DOI: 10.1136/emmermed-2012-201312
- Cooper, S., Kinsman, L., Buykx, P., McConnell-Henry, T., Endacott, R., Scholes, J., 2010. Managing the deteriorating patient in a simulated environment: Nursing students' knowledge, skill and situation awareness. *Journal of Clinical Nursing* 19 (15-16), 2309-2318. DOI: 10.1111/j.1365-2702.2009.03164.x
- Cooper, S., McConnell-Henry, T., Cant, R., Porter, J., Missen, K., Kinsman, L., Endacott, R., Scholes, J., 2011. Managing deteriorating patients: Registered nurses' performance in a simulated setting. *The Open Nursing Journal* 5 (1), 120-126. DOI: 10.2174/18744346011050100120
- Crocker, L., Algina, J., 2008. *Introduction to classical & modern test theory*. Cengage Learning, Mason, OH.
- Department of Health, 2009. *Competencies for recognising and responding to acutely ill patients in hospital*. Her Majesty's Stationery Office, London, UK.
- Ebel, R.L., 1965. *Measuring educational achievement*. Prentice-Hall, Englewood Cliffs, NJ.
- Emergency Nurses Association, 2007. *Trauma nursing core course (TNCC) provider manual*, 6th ed. Emergency Nurses Association, Des Plaines, IL.
- Endacott, R., Scholes, J., Buykx, P., Cooper, S., Kinsman, L., McConnell-Henry, T., 2010. Final-year nursing students' ability to assess, detect and act on clinical cues of deterioration in a simulated environment. *Journal of Advanced Nursing* 66 (12), 2722-2731. DOI: 10.1111/j.1365-2648.2010.05417.x
- Endsley, M.R., 1995. Measurement of situation awareness in dynamic systems. *Human Factors* 37(1), 65-84. DOI: 10.1518/001872095779049499

Endsley, M.R., 2000a. Direct measurement of situation awareness: Validity and use of SAGAT, in: Endsley, M.R., Garland, D.J. (Eds.), *Situation awareness analysis and measurement*. Lawrence Erlbaum, Mahwah, NJ, pp. 147-173.

Endsley, M.R., 2000b. Theoretical underpinning of situation awareness: A critical review, in: Endsley, M.R., Garland, D.J. (Eds.), *Situation awareness analysis and measurement*. Lawrence Erlbaum, Mahwah, NJ, pp. 3-32.

Fisher, D., King, L., 2013. An integrative literature review on preparing nursing students through simulation to recognize and respond to the deteriorating patient. *Journal of Advanced Nursing* 69 (11), 2375-2388. DOI: 10.1111/jan.12174

Gazarian, P.K., Henneman, E.A., Chandler, G.E., 2010. Nurse decision making in the prearrest period. *Clinical nursing research* 19 (1), 21-37. DOI: 10.1177/1054773809353161

Graham, J.M., 2006. Congeneric and (essentially) tau-equivalent estimates of score reliability: What they are and how to use them. *Educational and Psychological Measurement* 66 (6), 930-944. DOI: 10.1177/0013164406288165

Grove, S.K., Burns, N., Gray, J.R., 2013. *The practice of nursing research: Appraisal, synthesis, and generation of evidence*, 7 ed. Saunders Elsevier, St-Louis, MO.

Hansel, M., Winkelmann, A.M., Hardt, F., Gijsselaers, W., Hacker, W., Stiehl, M., Koch, T., Muller, M.P., 2012. Impact of simulator training and crew resource management training on final-year medical students' performance in sepsis resuscitation: A randomized trial. *Minerva Anestesiologica* 78 (8), 901-909.

Hoffman, K.A., Aitken, L.M., Duffield, C., 2009. A comparison of novice and expert nurses' cue collection during clinical decision-making: Verbal protocol analysis. *International Journal of Nursing Studies* 46 (10), 1335-1344. DOI: 10.1016/j.ijnurstu.2009.04.001

Hogan, J., 2006. Why don't nurses monitor respiratory rates of patients? *British Journal of Nursing* 15 (9), 489-492.

Hogan, M.P., Pace, D.E., Hapgood, J., Boone, D.C., 2006. Use of human patient simulation and the Situation Awareness Global Assessment Technique in practical trauma skills assessment. *Journal of Trauma - Injury, Infection and Critical Care* 61 (5), 1047-1052.

Laveault, D., Grégoire, J., 2014. *Introduction aux théories des tests en psychologie et en sciences de l'éducation*, 3rd ed. De Boeck Supérieur, Louvain-la-Neuve.

Liaw, S.Y., Scherpbier, A., Klainin-Yobas, P., Rethans, J.J., 2011. A review of education strategies to improve nurses' roles in recognizing and responding to deteriorating patients. *International Nursing Review* 58 (3), 296-303. DOI: 10.1111/j.1466-7657.2011.00915.x

Ludikhuizen, J., Smorenburg, S.M., de Rooij, S.E., de Jonge, E., 2012. Identification of deteriorating patients on general wards: Measurement of vital parameters and potential effectiveness of the

Modified Early Warning Score. *Journal of Critical Care* 27 (4), 424.e427-424.e413. DOI: 10.1016/j.jcrc.2012.01.003

McKenna, L., Missen, K., Cooper, S., Bogossian, F., Bucknall, T., Cant, R., 2014. Situation awareness in undergraduate nursing students managing simulated patient deterioration. *Nurse Education Today* 34 (6), e27-31. DOI: 10.1016/j.nedt.2013.12.013

NANDA International, 2012. *NANDA International nursing diagnoses: Definitions & classifications 2012-2014*. Wiley-Blackwell, Chichester, UK.

National Institute for Health and Clinical Excellence, 2007. *Acutely ill patients in hospital: recognition of and response to acute illness in hospital*. National Institute for Health and Clinical Excellence, London, UK.

O'Neill, E.S., Dluhy, N.M., Chin, E., 2005. Modelling novice clinical reasoning for a computerized decision support system. *Journal of Advanced Nursing* 49 (1), 68-77. DOI: 10.1111/j.1365-2648.2004.03265.x

Odell, M., Victor, C., Oliver, D., 2009. Nurses' role in detecting deterioration in ward patients: systematic literature review. *Journal of Advanced Nursing* 65 (10), 1992-2006. DOI: 10.1111/j.1365-2648.2009.05109.x

Pantazopoulos, I., Tsoni, A., Kouskouni, E., Papadimitriou, L., Johnson, E.O., Xanthos, T., 2012. Factors influencing nurses' decisions to activate medical emergency teams. *Journal of Clinical Nursing* 21 (17-18), 2668-2678. DOI: 10.1111/j.1365-2702.2012.04080.x

Polit, D.F., Beck, C.T., Owen, S.V., 2007. Focus on research methods: Is the CVI an acceptable indicator of content validity? Appraisal and recommendations. *Research in Nursing & Health* 30 (4), 459-467. DOI: 10.1002/nur.20199

Royal College of Physicians, 2012. *National Early Warning Score (NEWS): Standardising the assessment of acute-illness severity in the NHS*. Royal College of Physicians, London.

Sitterding, M.C., Broome, M.E., Everett, L.Q., Ebright, P., 2012. Understanding situation awareness in nursing work: A hybrid concept analysis. *Advances in Nursing Science* 35 (1), 77-92. DOI: 10.1097/ANS.0b013e3182450158

Streiner, D.L., Norman, G.R., 2008. *Health measurement scales: A practical guide to their development and use*. Oxford University Press, New York, NY.

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

Tanner, C.A., 2011. The critical state of measurement in nursing education research. *Journal of Nursing Education* 50 (9), 491-492. DOI: 10.1111/j.1365-2648.2006.03810.x\n10.3928/01484834-20110819-01

*Table 1*  
Socio-demographic data of the content validation experts

	Phase one ( <i>n</i> = 5)		Phase two ( <i>n</i> = 2)		Both phases ( <i>n</i> = 6)	
	M	SD	M	SD	M	SD
Age	39.20	7.16	33.00	9.90	37.00	10.47
Years of experience						
In profession	16.80	8.70	17.14	10.70	15.50	10.71
In critical care	11.60	8.02	12.67	9.62	11.33	8.19
	<i>n</i>		<i>n</i>		<i>n</i>	
Profession						
Advanced-practice nurse	5		1		6	
Senior emergency medicine resident	0		1		0	
Area of expertise						
Emergency care	2		1		2	
Intensive care	3		1		4	
Highest degree earned						
Baccalaureate	2		0		0	
Master	1		0		6	
Doctorate	2		2		0	
Major work domain						
Education	3		1		3	
Clinical practice	2		1		3	
Research activities	4		2		5	



Table 2

Full instrument with correct answer criteria

	Query <sup>1</sup>	Correct answer criteria
1-Perception (objective signs)	1 At the moment, what is the blood pressure?	±5 mmHg for systolic and diastolic
	2 At the moment, what is the heart rate?	±5 beats per minute
	3 At the moment, what is the respiratory rate?	±2 breaths per minute
	4 At the moment, what is the oxygen saturation?	±2% <sup>2</sup>
	5 At the moment, what is the level of consciousness? (AVPU)	
	6 At the moment, what is the patient's temperature?	±0.5°C <sup>3</sup>
1-Perception (less quantifiable signs)	7 At the moment, are his breath sounds normal?	
	8 At the moment, is his pulse regular?	
	9 At the moment, does he have difficulty breathing?	
	10 At the moment, does he need more oxygen?	
	11 At the moment, is he agitated?	
	12 At the moment, is he reporting unusual pain?	
	13 At the moment, is he reporting increasing pain?	
	14 At the moment, is he reporting that something serious is about to happen to him?	
2-Comprehension	15 Do you think his airway is patent?	
	16 Do you think his respiration is efficient?	
	17 Do you think his cardiac output is normal?	
	18 Do you think his peripheral perfusion is normal?	
	19 Do you think he is hypothermic or hyperthermic?	
	20 Is he showing signs of shock?	
	21 Is he showing signs of neurological involvement?	
	22 Is he showing signs of internal or external bleeding?	
	23 Is he showing signs of infection?	
3-Projection	24 In the next few minutes, will you have to administer a bolus? <sup>4</sup>	
	25 In the next few minutes, will you advise the doctor of your observations?	
	26 In the next few minutes, will you ask the doctor to come STAT <sup>5</sup> to the patient's bedside? <sup>4</sup>	
	27 In the next few minutes, what will happen to his blood pressure?	
	28 In the next few minutes, what will happen to his heart rate?	
	29 In the next few minutes, what will happen to his respiratory rate?	
	30 In the next few minutes, what will happen to his oxygen saturation?	
	31 In the next few minutes, what will happen to his systemic circulation?	

NOTE: <sup>1</sup>Maximum score for each query = 1 point, except for query 1 (1 point for systolic and 1 point for diastolic blood pressure, maximum score = 2 points). <sup>2</sup>If the manikin's SpO<sub>2</sub> is 94–100%, the student's answer must be in this range to be correct. If the SpO<sub>2</sub> < 94%, the student's answer must be in that range to be correct (given that a SpO<sub>2</sub> value ≥ 94% is considered normal in the absence of a pulmonary condition). <sup>3</sup>If the manikin's T° is between 36.0–37.5°C, the student's answer must be in that range to be correct. If the manikin's T° is outside the normality range (36°C > T° > 37.5°C), the student's answer must be outside that range to be correct. <sup>4</sup>Exploratory items. <sup>5</sup>From Latin "statim," that is, immediately.

*Table 3*  
Socio-demographic data of participating students ( $n = 234$ )

	First cohort ( $n = 109$ )		Second cohort ( $n = 77$ )		Third cohort ( $n = 48$ )	
	M	SD	M	SD	M	SD
Age	23.57	3.88	25.20	7.52	23.10	3.97
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%
Gender						
Male	14	12.84	11	14.29	5	10.42
Female	95	87.16	66	85.71	43	89.58
Program of study						
Entry-to-practice	95	87.16	56	72.73	29	60.42
Post-diploma	14	12.84	21	27.27	19	39.58
Work experience in nursing						
None	98	89.91	60	77.92	31	64.58
< 1 year	2	1.83	4	5.19	5	10.42
1–2 years	6	5.50	11	14.29	11	22.92
3–4 years	3	2.75	1	1.30	1	2.08
> 4 years	0	0.00	1	1.30	0	0.00

*Table 4*  
Results of item analysis

Query	<i>p</i>	<i>D</i>	$\rho$ pbis-i
Level one (M = 9.44, SD = 2.60)			
Objective signs			
1.1 – Systolic blood pressure	0.73	0.69	0.51
1.2 – Diastolic blood pressure	0.47	0.76	0.43
2 – Heart rate	0.57	0.78	0.47
3 – Respiratory rate	0.24	0.41	0.20
4 – Oxygen saturation	0.62	0.66	0.36
5 – Level of consciousness	0.86	0.18	<b>0.01</b>
6 – Temperature	0.44	0.66	0.32
Less quantifiable signs			
7 – Breath sounds	0.50	0.56	0.20
8 – Pulse regularity	0.66	0.58	0.25
9 – Breathing difficulty	0.73	0.56	0.30
10 – Oxygen needs	0.63	0.51	0.19
11 – Agitation	0.89	0.18	0.14
12 – Unusual pain	0.84	0.16	<b>0.01</b>
13 – Increasing pain	0.58	0.48	<b>0.10</b>
14 – Something serious	0.68	0.68	0.32
Level two (M = 6.01, SD = 1.61)			
15 – Airway patency	0.92	0.24	0.30
16 – Respiration efficiency	0.64	0.55	0.20
17 – Cardiac output normality	0.33	0.46	0.16
18 – Peripheral perfusion normality	0.71	0.07	<b>-0.17</b>
19 – Hypo/hyperthermia	0.55	0.70	0.25
20 – Shock	0.60	0.62	0.29
21 – Neurological involvement	0.90	0.23	0.21
22 – Bleeding	0.54	0.27	<b>-0.03</b>
23 – Infection	0.84	0.44	0.34
Level three (M = 6.18, SD = 1.66)			
24 – Bolus administration	0.55	0.76	0.29
25 – Advise doctor	0.91	0.18	0.15
26 – Doctor at bedside	0.46	0.35	<b>-0.18</b>
27 – Projection blood pressure	0.81	0.55	0.61
28 – Projection heart rate	0.89	0.31	0.49
29 – Projection respiratory rate	0.86	0.41	0.50
30 – Projection oxygen saturation	0.86	0.45	0.62
31 – Projection systemic perfusion	0.86	0.39	0.49

NOTE: Insufficient  $\rho$  pbis-i are marked in bold

*Figure 1*  
 Relationship between difficulty and discrimination indices

