education sciences

MDPI

Article

Score Your Way to Clinical Reasoning Excellence: SCALENEo Online Serious Game in Physiotherapy Education

Renaud Hage [1,2,3,*], Frédéric Dierick [1,3,4], Joël Da Natividade [4], Simon Daniau [5], Wesley Estievenart [1], Sébastien Leteneur [6], Jean-Christophe Servotte [5], Mark A. Jones [7,8] and Fabien Buisseret [1]

1 Laboratoire Forme et Fonctionnement Humain, Haute Ecole Louvain en Hainaut (HELHa), rue Trieu Kaisin 136, 6061 Montignies-sur-Sambre, Belgium; frederic.dierick@rehazenter.lu (F.D.); estievenartw@helha.be (W.E.); buisseretf@helha.be (F.B.)
2 Traitement Formation Thérapie Manuelle (TFTM), Manual Therapy Center, Avenue des Cerisiers 211A, 1200 Brussels, Belgium
3 Faculté des Sciences de La Motricité, UCLouvain, Place Pierre de Coubertin 1-2, 1348 Louvain-la-Neuve, Belgium
4 RehaLAB, Centre National de Rééducation Fonctionnelle et de Réadaptation—Rehazenter, Rue André Vésale 1, 2674 Luxembourg, Luxembourg; joel.danatividade@rehazenter.lu
5 Fors Recherche & Développement, Haute école de Namur-Liège-Luxembourg (Henallux), Rue Joseph Calozet 19, 5000 Namur, Belgium; simon.daniau@henallux.be (S.D.); jean-christophe.servotte@henallux.be (J.-C.S.)
6 LAMIH CNRS UMR 8201, Université Polytechnique Hauts-de-France, F-59313 Valenciennes, France; sebastien.leteneur@univ-valenciennes.fr
7 Allied Health and Human Performance Academic Unit, University of South Australia, Adelaide 5000, Australia
8 International Centre for Allied Health Evidence, University of South Australia, Adelaide 5000, Australia
* Correspondence: renaud.hage@uclouvain.be

Abstract

SCALENEo (Smart ClinicAL rEasoning iN physiothErapy) is an innovative online serious game designed to improve clinical reasoning in musculoskeletal physiotherapy education. Adapted from the "Happy Families" card game, it provides an interactive, structured approach to developing students/learners' ability to categorize clinical information into families of hypotheses. This digital platform supports both self-directed and collaborative learning, eliminating the need for continuous instructor supervision while ensuring meaningful engagement. SCALENEo features a unique feedback and scoring system that not only assesses students/learners' decision-making processes but also promotes cautious and reflective reasoning over random guessing. By aligning with evidence-based pedagogical strategies, such as serious games and formative assessment, SCALENEo offers educators a powerful tool to reinforce critical thinking, improve student/learner engagement, and facilitate deeper learning in clinical reasoning education.

Keywords: clinical reasoning; serious games; physiotherapy education; digital learning

check for updates

Academic Editor: Diego Vergara

Received: 22 July 2025
Revised: 17 August 2025
Accepted: 18 August 2025
Published: 21 August 2025

Citation: Hage, R., Dierick, F., Da Natividade, J., Daniau, S., Estievenart, W., Leteneur, S., Servotte, J.-C., Jones, M. A., & Buisseret, F. (2025). Score Your Way to Clinical Reasoning Excellence: SCALENEo Online Serious Game in Physiotherapy Education. *Education Sciences*, *15*(8), 1077. https://doi.org/10.3390/educsci15081077

1. Introduction

Learning clinical reasoning (CR) is an integral component of physiotherapy education (Schuwirth, 2009) and professional practice, directly influencing decision-making processes crucial for effective and safe patient care. CR encompasses a range of analytical and non-analytical processes and skills used in patient assessment and management, involving both the development of hypotheses and pattern recognition (Croskerry, 2009a, 2009b; Jones, 1992). For physiotherapy students, mastering hypothetico-deductive analytical reasoning—involving generating and testing multiple hypotheses—is essential due to the variability and complexity of clinical presentations (Bowen, 2006; Jones, 1992; Jones et al., 2019; Jones & Rivett, 2004).

The importance of employing different categories of hypotheses (Jones et al., 2019), also known as families of hypotheses (Hage et al., 2023), is well-documented in the reflective practices of experienced clinicians and educators (Barlow, 2012; Edwards et al., 2004; Jones et al., 2019; Rushton et al., 2016; Smart & Doody, 2007). This structured approach enhances clinical judgment consistency, fosters thorough analysis of patients' cases, and mitigates cognitive biases. However, despite their recognized value, teaching and assessing CR remain challenging. Traditional methods, such as case-based discussions or direct supervision, are time-consuming and may not always provide interactive and adaptive learning opportunities suited to diverse student needs.

In response to these challenges, a variety of tools and methods are needed in educational settings to optimize learning for the diverse profiles and skills students possess. One promising approach is serious games (SGs), a type of game-based learning aimed at achieving specific educational outcomes (Gorbanev et al., 2018; Haoran et al., 2019), which have shown efficacy in improving concentration and knowledge retention among learners (García-Redondo et al., 2019; Gorbanev et al., 2018; Tubelo et al., 2019). A growing body of literature supports the integration of SG into health education as a solution to this challenge (Blanié et al., 2020; Gorbanev et al., 2018; García-Redondo et al., 2019). SGs are digital or physical games designed with an explicit educational goal, offering structured, engaging, and repeatable learning experiences (Haoran et al., 2019). CR learning through SGs (Blanié et al., 2020; Koivisto et al., 2018) and online simulation games (Cant & Cooper, 2010; Lapkin et al., 2010) has already been successfully implemented in nursing education, showing promising results in improving hypothesis-driven analysis and structured decision-making. SG and online simulation tools have been successfully applied in nursing education, where they have demonstrated clear benefits in developing hypothesis-based reasoning and structured decision-making skills (Blanié et al., 2020; Koivisto et al., 2018; Cant & Cooper, 2010; Lapkin et al., 2010).

The educational value of SGs has since been extended to other health professions, including medicine, with evidence supporting improvements in engagement, knowledge retention, and CR (Tubelo et al., 2019; Qiao et al., 2023). Inspired by these successes, we extended the use of SG learning to physiotherapy education through the "Happy Families" card game (Hage et al., 2023). The latter game is a kind of clinical-case-analysis gamification. Gamification refers to applying game concepts (e.g., points, progress, and narrative) layered on non-game learning contexts to make learning more enjoyable and engaging; these principles provide guidance for designing educational tools that enhance engagement and learning outcomes (Connolly et al., 2012; Koivisto et al., 2018; Krishnamurthy et al., 2022).

The "Happy Families" SG engages students in learning CR by raising awareness of cognitive biases and promoting reflective practice, mainly through the perception of relevant or significant information when analyzing a clinical case using hypothetico-deductive reasoning. Moreover, because the social and cultural environment in learning is important, it also takes place in a social framework that is supported by socio-constructivist learning theory (Clark & Dumas, 2015; Hickey, 1997). In fact, learning and the motivation to learn are reinforced by exchanges within groups of learners. At the start of the game, players (i.e., students, typically in groups of 3 to 5 players with a maximum of 5 groups per session) are provided with a fully detailed clinical case including both quantitative and qualitative patient data. The game includes a set of cards representing the hypothesis families: core categories of clinical judgment such as patients' expectations, pain, origin of symptoms, etc.; see Hage et al. (2023) for a full explanation. Students must then select the cards where they have initial hypotheses and, using an erasable pen, write a list of their hypotheses on the back of the card. In addition, a "wild card" is provided for players to write their main hypotheses (i.e., first choice at this stage) in each family. The game advances as players categorize hypotheses and refine their reasoning. The game concludes once all hypotheses have been categorized. Then each group presents its analysis to the other groups, allowing them to discuss, justify, and compare their hypotheses. Interaction between the players and feedback from the teacher helps to refine or modify the hypotheses. Because structured, timely feedback helps learners identify information gaps, refine hypotheses, and consolidate CR, we prioritized embedding multi-step formative feedback within the Happy Families digital platform.

Beyond SG, technology-enhanced learning is increasingly recognized as a vital tool for improving both education and continuing professional development, which are essential for ensuring high-quality healthcare services. The positive impact of using virtual scenarios as learning activities on outcomes such as knowledge acquisition and CR has been well-established in medicine (Kononowicz et al., 2019). These technological methods can be described as "interactive computer simulations based on clinical scenario cases". In a 2019 systematic review and meta-analysis assessing the effectiveness of virtual patients compared to traditional teaching, Kononowicz et al. (2019) found that using virtual patients improved or equaled traditional education in terms of CR, procedural skills, and a combination of procedural and team skills.

The Erasmus+ "SCALENEo" (Smart ClinicAL rEasoning iN physiothErapy) project (https://scaleneo.eu) (accessed on 19 august 2025) aims to promote the SG Happy Families through an online approach. The aim is to leverage the possibilities offered by information and communication technologies to develop an online version of this SG to reach geographically remote and emerging audiences and, mostly, to allow individual players, i.e., students or physiotherapists, to participate in ongoing professional development of CR. While maintaining the same theoretical concepts, online gaming operates differently from the face-to-face game. In the face-to-face game form, participants must list the information found in the clinical case, related to each hypothesis family, so the same information can be found in different families of hypotheses (Figure 1A). In the online version, the sentences explaining the clinical case can be linked to the hypothesis families (Figure 1B). The online implementation of SG leads to a key question that is at the heart of the present study: How to provide relevant feedback about one player's CR? As in test-enhanced learning (Roediger & Karpicke, 2006), it is indeed crucial to evaluate students and to provide feedback on the quality of their work and their progress in learning CR. Thanks to the evaluation, the Happy Families game can, therefore, be used as a formative evaluation tool, provided that the students can be evaluated.

The value of student assessment is well-established (Larsen et al., 2008). CR skills are particularly challenging to assess because they encompass a combination of cognitive abilities required to analyze and synthesize problems (Min Simpkins et al., 2019). Research in cognitive psychology indicates that the use of tests can enhance understanding and application (Larsen et al., 2008). Evaluating CR, which stems from complex internal cognitive processes that are not directly observable, is inherently problematic. CR involves the integration of cognitive, psychomotor, and affective skills (Huhn et al., 2019), necessitating a flexible assessment procedure. The term CR encompasses various skills including decision-making, critical thinking, problem-solving, clinical judgment, and diagnostic reasoning (Young et al., 2019). These terms can help identify and assess the relevant skills. Combined, they contribute to a global assessment.

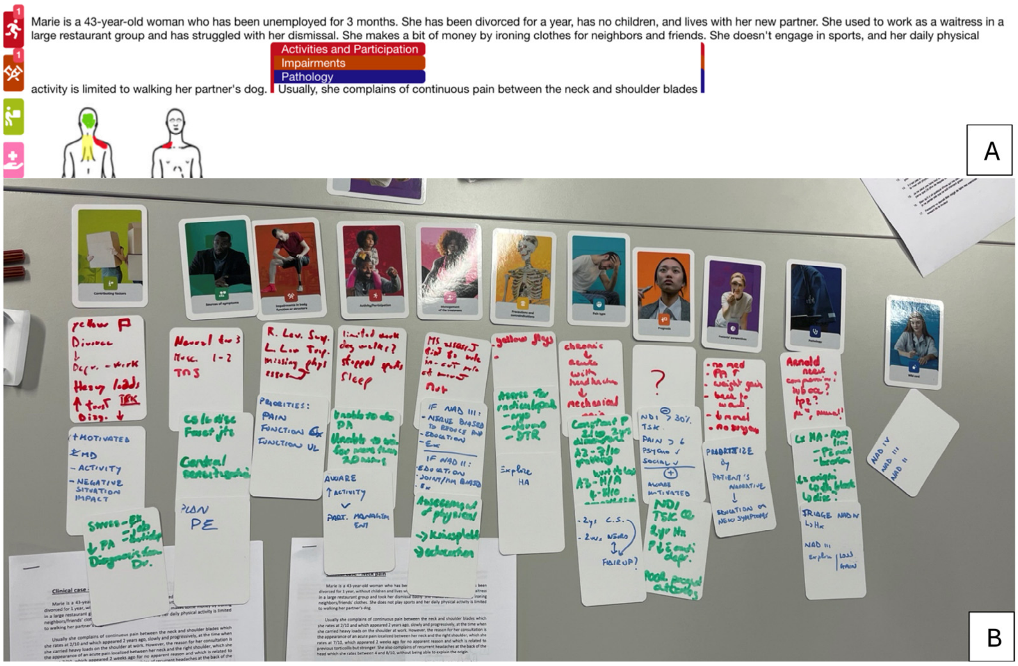


Figure 1. (**A**) Online version of the SG where participants assign statements related to the clinical case to one or more hypothesis families. (**B**) In the face-to-face version, participants write down their hypotheses by hand and organize them into families. This version displays all the families alongside the hypotheses derived from the clinical case, as provided by the participants. Note: The images are intended solely to illustrate the SG in its different formats. The answers shown are for illustrative purposes only and do not necessarily represent correct responses.

Despite the demonstrated benefits of SG in the nursing and medicine fields, there is a notable lack of research on the use of such tools for assessing and enhancing CR in physiotherapy. Moreover, in the medical field, there is no gold standard for assessing CR (Min Simpkins et al., 2019), and this is also true in physiotherapy, highlighting a significant gap in the literature. Current assessment methods in medical education, such as the Assessment of Reasoning Tool (Thammasitboon et al., 2018), the Think Aloud protocols (Pinnock et al., 2015), and the Script Concordance Test (SCT) (Charlin et al., 2000) have been proposed. Specifically, the SCT method has been used since the work of Charlin et al. (2000). The SCT is based on script theory and involves comparing students' interpretation of clinical data with those of experienced clinicians through a series of clinical tasks presented in specific and "ill-defined" contexts (Charlin et al., 2006). In summary, an "ill-defined" case is presented to candidates as a brief clinical scenario where the information provided is insufficient to make a definitive decision. In addition, data provided by motion capture systems is integrated into the scenario to improve their integration in the participants' initial CR processes. Thus, biomechanical data (e.g., kinematic values) that illustrate the clinical case are presented in connection with the clinical case. The objective is to enrich the participant's reasoning with quantitative, sensor-based data. Each case is followed by several questions that begin with a hypothesis, followed by additional information. Participants are then asked to assess how this added information impacts the likelihood of the proposed hypothesis being correct (Brentnall et al., 2022; Charlin et al., 2000; Dory et al., 2012, 2016).

Our proposal, adapting SCT to our context, seeks to create an engaging and context-rich environment for assessing reasoning under uncertainty, which is a hallmark of real-life clinical decision-making. Given that SCT is rooted in the evaluation of reasoning in ambiguous situations, the potential advantages and limitations of such an assessment format, particularly in a gamified setting, deserve further exploration. Central to this method is the aggregate scoring approach, which has been consistently endorsed in the literature (Charlin et al., 2000; Daniel et al., 2019; Deschênes et al., 2011; Dory et al., 2012; Lubarsky et al., 2011). This scoring method emphasizes the relationship between the learner's selection of answers most frequently chosen by experts. In other words, SCALENEo captures concordance in information triage (which items are attended to and how they are categorized) and hypothesis weighting (probability judgments) relative to experts, but it does not operationalize the wider spectrum of CR competencies, including the full scope of diagnostic reasoning and management decision-making.

In the original SCT, the information gradually added by the person who prepared the clinical case must be confirmed or refuted to strengthen a hypothesis. In SCALENEo, the players themselves must add the additional clinical information present in the clinical case to reinforce or refute an existing family of hypotheses. At the end of the game, the players are asked to evaluate their choices based on the probability that the proposed hypotheses are correct.

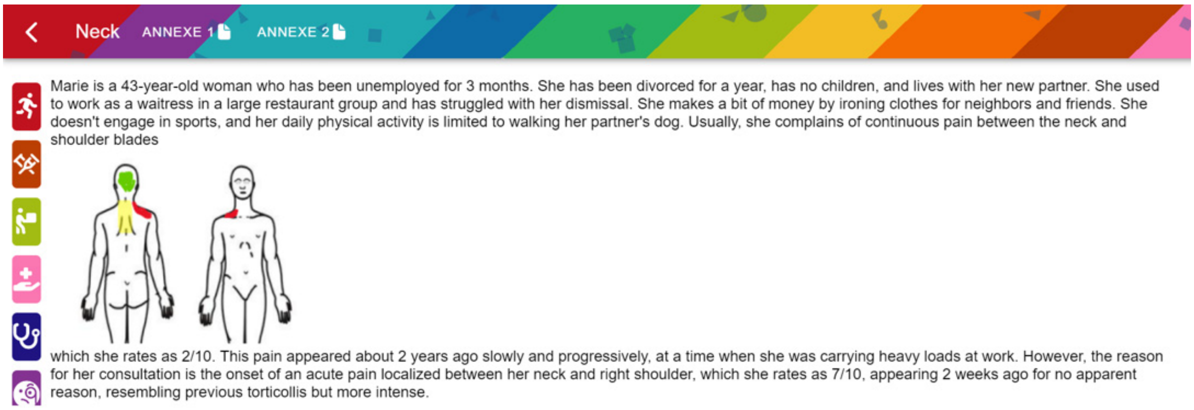
Since the SCT method has proved reliable and effective for assessing CR in health professions such as nursing (Deschênes et al., 2011), medicine (Cohen Aubart et al., 2021), and physiotherapy (Kojich et al., 2024; O'Brien et al., 2023), the aim of this study was to examine the feasibility of adapting an SCT-inspired scoring approach within the online Happy Families game to assess CR in physiotherapy. The research question was, therefore, as follows: To what extent does the adapted SCT scoring differentiate performance (i.e., agreement with expert item categorization) across simulated error patterns, and can it support formative feedback in training contexts? To answer our research question, our article will be organized as follows. First, the methodology will be explained, including a description of the Happy Families game and of its online implementation. Next, the scoring algorithm will be explained in detail, and the results will then be presented and will consist of a computation of how our score evolves with the number of errors a player makes. Finally, we will conclude with a discussion regarding the relevance of the online SCALENEo application and of its scoring and feedback algorithm.

2. Materials and Methods

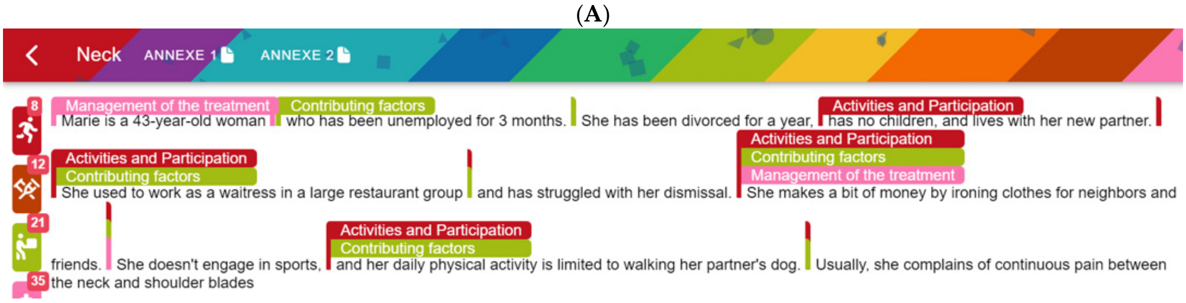
First, a description of how to play the game will be given (https://scaleneo.eu/en/home-english/ (accessed on 19 August 2025)). Then, the technical details needed to understand the feedback will be provided.

2.1. Online Game

The game starts when a player or a group of players selects a given clinical case, the clinical cases being labeled by pain localization (e.g., neck pain, hip pain, etc.). The player(s) then read(s) the clinical case, mostly composed of text and biomechanical data. Pre-selected groups of words or elements can be, by clicking, selected and assigned to one or more hypothesis families in the form of a card (Figure 2A). The player can select an element of information (an item) by clicking once to mark the item as relevant or clicking twice to mark it as significant, and then assigning it to a hypothesis family card. "Significant" means an element that players must identify and highlight as critical. This first phase of the game, therefore, requires hypothetico-deductive reflection on the clinical case and the selection and linking elements of information to the hypothesis family or families the player judges relevant. It must be said that for each clinical case, a reference game exists, i.e., a "perfect" game obtained by consensus by the authors of the present paper. In this reference game, the authors have not only classified relevant information into the appropriate hypothesis families but also labeled the items by importance.



(**A**)



(**B**)

Figure 2. Screenshots of a neck pain clinical case. The clinical case is displayed before assignments of the information to hypothesis families (**A**) and after (**B**). Note: The assignments in panel B are not necessarily correct and are taken from a random group of players' games.

During the game, to foster interaction, a live chat is available during multiplayer games, enabling players to discuss reasoning strategies in real-time. In classroom settings, students are also encouraged to interact with their instructor during gameplay to clarify doubts or justify choices. While digital interaction cannot fully replicate face-to-face dialogue, the SCALENEo platform strives to approximate collaborative CR through structured peer discussion and feedback mechanisms. Players are allowed to consult their courses or any relevant source of information—articles published in peer-reviewed journals are also suggested in the clinical cases.

Figure 3 illustrates the feedback process in SCALENEo. After completing their analysis (Figure 3A), players receive hypothesis-by-hypothesis feedback using a colored indicator system (Figure 3B). Green indicates that all significant elements are found and more than 50% of the reference card elements are correctly identified, yellow indicates that all significant elements are found and the percentage of correctly identified items is below the threshold of 50, and red indicates that significant elements have been missed or classified in the wrong family card or incorrectly marked as relevant. A global appreciation is then displayed (Figure 3C), with brain icons (gold/silver/bronze) based on the number of attempts required to eliminate errors. This two-level feedback helps reinforce both local reasoning and overall strategy. Three difficulty levels are implemented, giving the player(s) the opportunity to make several attempts before receiving definitive feedback depending on the level of difficulty: easy is 3 attempts, medium is 2, and hard is 1. The global brain level score is calculated according to the number of attempts needed to eliminate all the red lights: first attempt, gold; second attempt, silver; and third attempt, bronze. If, after the allowed attempts, there are still cards with red lights, no brain level is awarded, which signifies the player(s)'s failure to identify and classify all the significant elements.

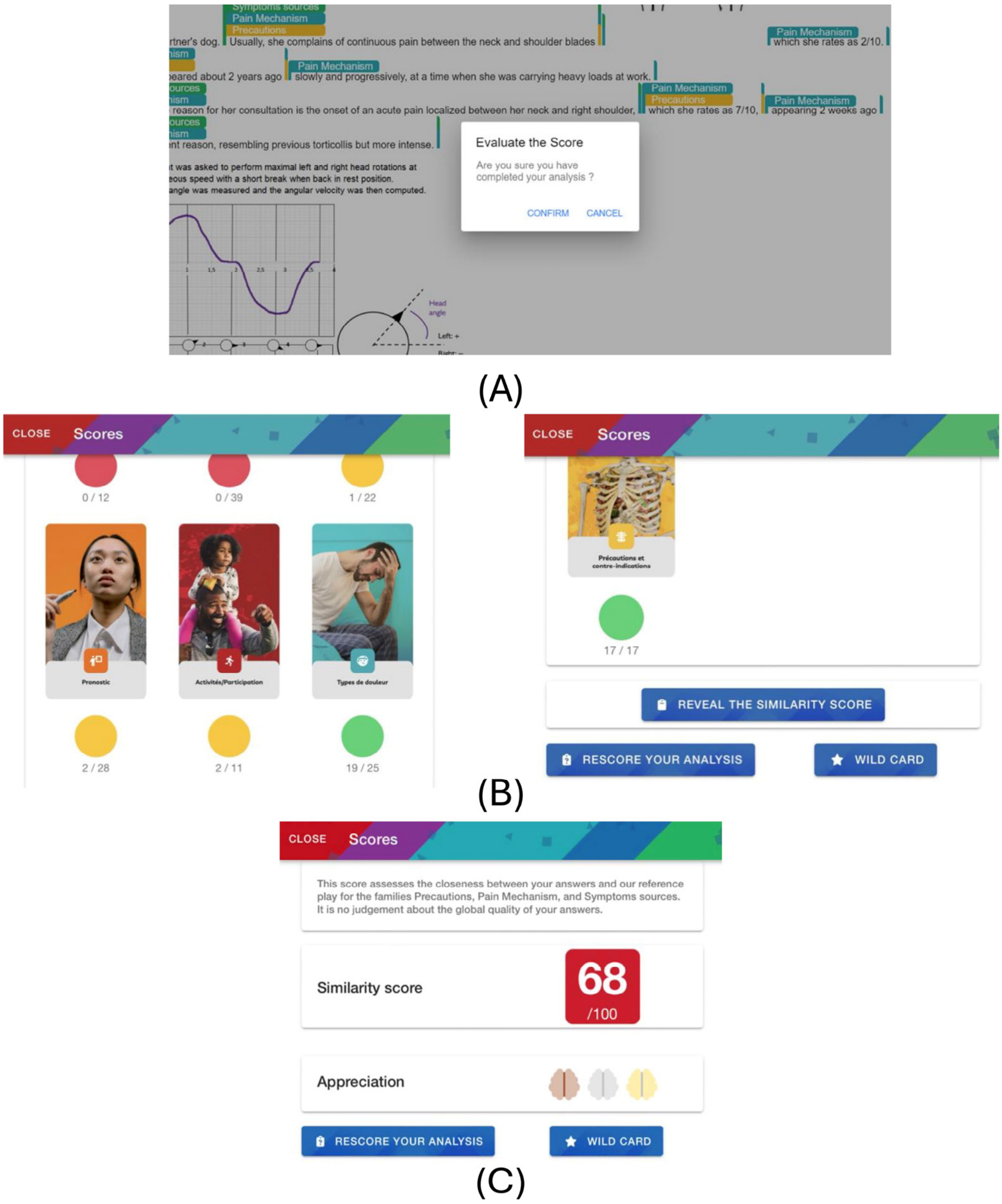


Figure 3. (**A**) End of a game and beginning of the assessment procedure. (**B**) Hypothesis-by-hypothesis assessment; counting of the number of elements correctly assigned by the player(s). (**C**) The score is revealed with a warning message and a global appreciation via gold/silver/bronze brain.

At the end of all attempts, a "wild card" must be selected (Figure 4A). This card represents, according to the player(s), the key information guiding the CR and physical management of the presented case. All hypothesis categories must be considered, and those related to precautions and contraindications may be prioritized when necessary. A final score is then computed. After the choice of the "wild card", the final feedback is displayed (Figure 4B): It consists of an overall score and of all the other elements mentioned above. At this stage it is also possible to look at the reference game. The total score is based on the average card score for the families of clinical judgment considered most important by the authors of this study. For example, if the hypotheses "type of pain", "sources of symptoms", and "precautions and contraindications" were judged to be most important to the selected clinical case, then the total score would be based on those families. The total score is updated after the last attempt and is now the average score between the 3 above families and the "wild card" score.

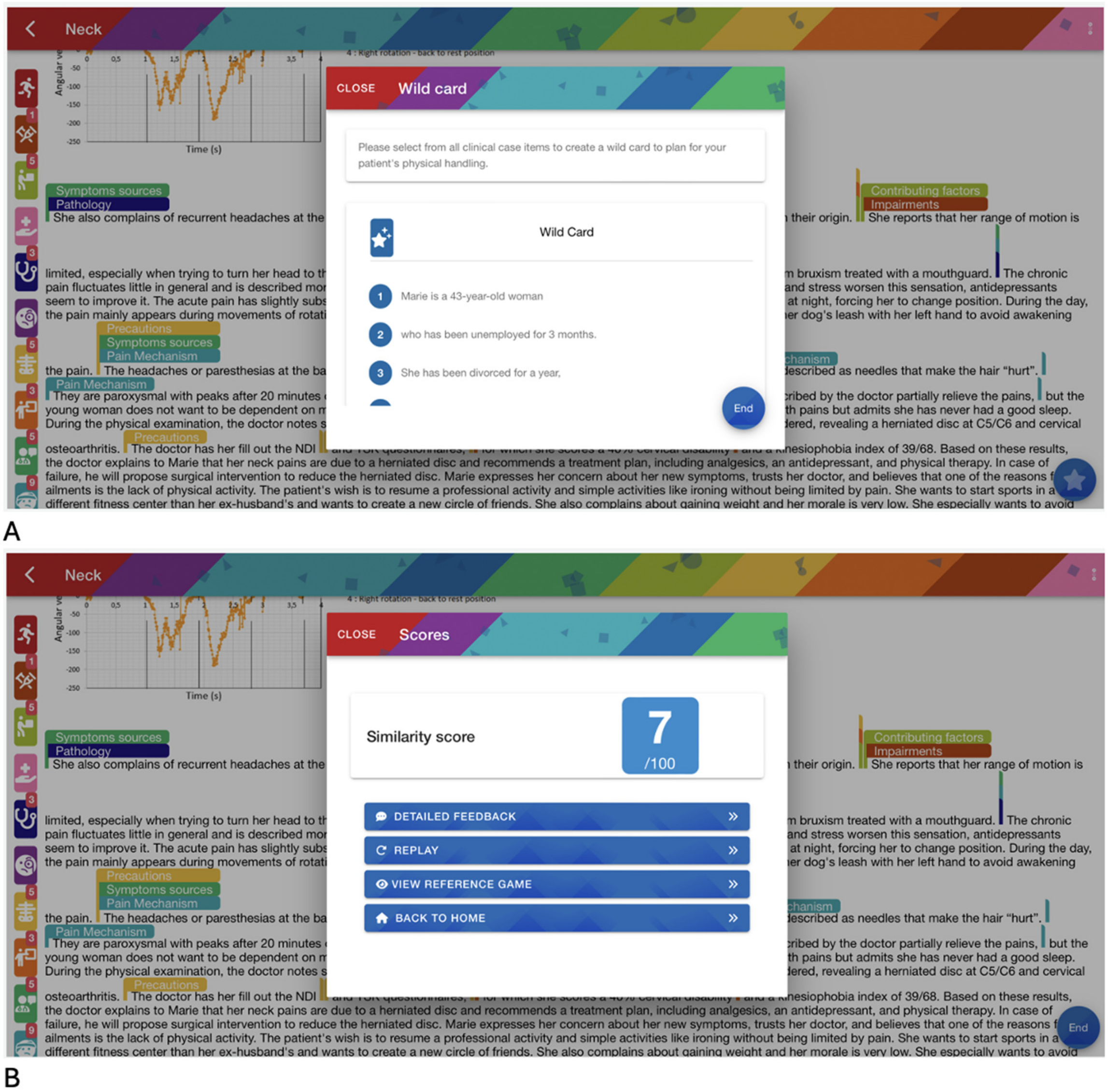


Figure 4. (**A**) Selection of the "wild card". (**B**) Final feedback. Note that the score is exceptionally low since the image is of a randomly generated game.

A first version of the feedback procedure has been tested by a panel of 6 third-year physiotherapy students at the University of Quebec at Chicoutimi. They reviewed the process, and their feedback led us to change the order of the different steps. In our first version, the score was displayed first, but it appeared to be meaningless to the students. The order presented in this paper, with the total score being shown at the end of the process, is the result of these early tests.

2.2. Classifying Card Elements

We now turn to the technical issues "hidden" behind the online game; these points are a key part of the novel approach we propose to assess CR.

Technically, a played game consists of a set of selected items by the player(s) and the game cards that must be compared to the reference cards, i.e., the cards filled by the SCALENEo experts. The comparison is made item by item and detailed in Table 1.

Table 1. Schematic view of one element's classification according to its presence or not in a game card. Relevant items are those present in the reference game but not considered significant. Significant items, on the other hand, are essential elements in the reference game that players must identify and highlight as critical. Good items are those that have been correctly assigned by the player(s) to the appropriate hypothesis family. Missed items refer to those present in the reference game but not identified or selected by the player(s). Wrong items are those selected by the player(s) but not present in the reference game. Bad items are those selected by the player(s) but incorrectly assigned to the wrong hypothesis family. Finally, empty items are those that are neither selected by the player(s) nor present in the reference game. Green color is used for good/bad items, orange for wrong items, purple for missed items, and gray for empty items.

		Relevant/Significant Item in Reference Card	
		Present	**Not Present**
Relevant/Significant Item in Game card	Present	good/bad	wrong
	Not present	missed	empty

After a comparison between the reference and game cards, the lights are displayed. Their color is as follows:

- Green if the game and reference cards are either both empty or if all the significant elements are found and more than 50% of the reference card elements are correctly identified. This threshold of 50% is a parameter and could be lowered or increased according to the player(s) experience, for example.
- Yellow if all the significant elements are found and the percentage of correctly identified items is below the threshold of 50%.
- Red if significant elements have been missed or classified in the wrong family card, or incorrectly marked as relevant.

2.3. Total Score

The total score is based on the average card score for the families of judgment considered most important by the authors of this study. To compute the card score, a Cohen's Kappa coefficient measuring the agreement between the played game and the reference one is computed—the game and reference cards are, therefore, seen as two independent observers whose agreement can be measured. Cohen's Kappa is finally multiplied by 100 to have a card score between 0 and 100. Note that if a card in the reference game is empty and the corresponding card in the game played is also empty, the score is defined as 100%, because the two observers agree not to put any elements in the cards. Figure 5 displays 3 examples of card score calculations.

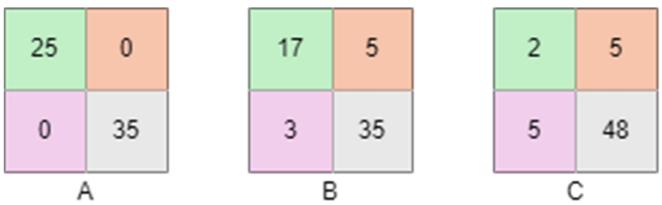


Figure 5. Examples of distribution of elements in a card according to the classification described in Table 1: (**A**) No wrong or missed items are present, leading to a score of 100%. (**B**) Some missed or wrong items are present; leading to a score of 71%. (**C**) A low number of good/bad items are present, leading to a score of 19%. Green color is used for good/bad items, orange for wrong items, purple for missed items, and gray for empty items.

2.4. Score Evolution vs. Number of Errors

The following procedure is proposed to assess our way of scoring a game: First, using a developer tool in the online application, game samples were generated by duplicating one clinical case's reference game (neck pain patient), and then by introducing errors by randomly removing a given percentage (n) of items. Twenty samples for each value of n were generated with n = 0, 5, 10, . . ., 100%. Second, the same procedure was applied, but items were randomly replaced by a random item instead of being removed. A total of 840 plays were generated. In both cases, at a given n, interpreted as the % of errors, the average score and its standard deviation were computed for a single attempt (no "wild card" choice has been performed). Then, the score decrease rate was computed as the derivative of the score as a function of n, where the latter function of n was obtained by a fourth-order regression line.

The games with removed items are interpreted as representative of a "cautious player", i.e., a player preferring not to answer when they are not certain of the answer. The games with replaced items are interpreted as representative of a "bold player", i.e., a player taking the risk of answering to items of which they are not fully sure.

3. Results

The evolution of the average score versus the percentage of error is displayed in Figure 6. The red line shows the generated plays with removed items, and the blue line shows the generated plays with replaced items. The fourth-order polynomial regression lines have the following equations: Score = $98.8 - 14.1\ n - 266\ n^2 + 254\ n^3 - 71.1\ n^4$ (removed items) and Score = $99.9 - 104\ n - 58.1\ n^2 + 65.3\ n^3 - 2.8\ n^4$ (replaced items), where we recall that n is the percentage of errors. It can be observed that the scores with replaced items are always lower than with removed items at a given percentage of error: The bold player always has a lower score than the cautious one in this approach. Thus, the scoring function discriminated error profiles, supporting feasibility of SCT-inspired scoring for formative CR assessment.

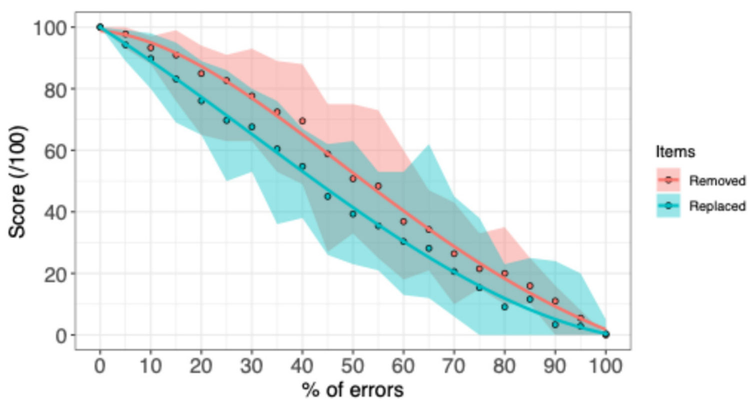


Figure 6. Evolution of the averaged scores computed on 20 plays versus the percentage of errors (%). The red and blue points denote games where items are randomly removed or replaced, respectively. The shaded areas define an interval of ± 1 standard deviation around the average scores. The solid lines are fourth-order polynomial regression lines.

The score decrease rate versus the percentage of error is displayed in Figure 7. The negative score decrease rate observed at a given percentage of error can be interpreted as the score variation corresponding to a unitary increase in percentage of error. In other words, to what extent "one more error" has a strong impact or not on the total score. It is observed that the score is not very sensitive to the % of error if the latter is small for a cautious player (score decrease rate close to zero in the case where items are removed).

Both cautious and bold payers' scores are strongly impacted by the % of error as soon as the latter is larger than about 25%: the score tends to decrease by about 1/60 for each new % of error. The numerical values are given in the Appendix A; see Tables A1 and A2 for completeness.

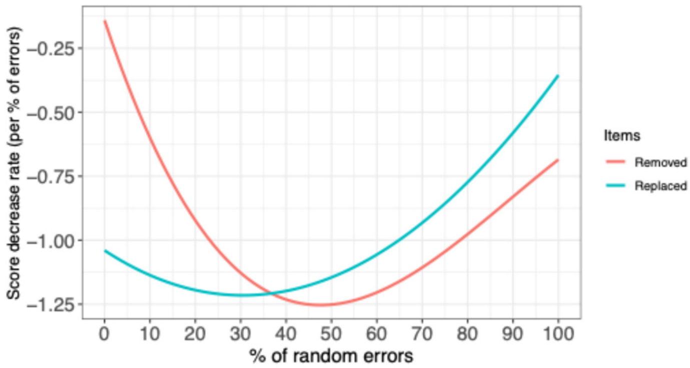


Figure 7. Same as Figure 6 for the score decrease rate, computed as the derivative of the score regression lines versus the percentage of random errors.

These results suggest that the integration of an SCT-based scoring method into SCALENEo can effectively discriminate between different CR strategies. The consistently lower scores observed in bold players, as compared to cautious ones, as well as the sensitivity of scores to increasing error rates, indicate that the scoring system captures variations in reasoning performance. This supports the feasibility of using such a game-based format for assessing CR under uncertainty. While the study does not directly measure improvements in learning outcomes, the ability to identify cognitive errors and provide score-based feedback suggests a strong potential for formative use in physiotherapy education.

4. Discussion

4.1. Feedback in CR Assessment

Our SG shares similarities with SCT, a well-established method for assessing CR. In SCT, participants assess a clinical situation using limited information, after which additional information is introduced to reassess their previous choices. These are then evaluated against expert judgments (Dory et al., 2016; Charlin et al., 2000). Conversely, in SCALENEo, players must identify and categorize information from a comprehensive clinical case presented upfront. This design fosters active engagement, encourages players to justify their thought processes, and strengthens their ability to navigate complex and ill-defined scenarios—a key aspect of effective CR (Croskerry, 2009b). In our clinical cases, each card can be considered as a "case" that will need to be categorized according to each family of hypotheses. That is why the first feedback we provide is the collection of lights, displayed card-by-card. This aims at showing the player(s) which hypotheses families they understand satisfactorily (or poorly or not at all), and, maybe most importantly, ensuring that the clinical red flags are clearly identified so students using this SG can prepare for the first line in clinical practice. Red flags are used to help clinicians identify serious pathology, and the majority of guidelines recommend the use of red flags (Finucane et al., 2020). One of SCALENEo's key innovations is its two-step feedback system: (1) card-by-card feedback; and (2) global scoring. We recall that the use of a predefined set of clinical hypotheses is one of the main features of the Happy Families game we developed. It allows players to engage in answer generation while keeping a structured framework in which players formulate their answers. Because players actively select and tag case data and must nominate a "wild card" key driver, the task requires generation and prioritization rather than simple recognition. However, a limitation of our framework is that we do not allow for the creation of a new family of hypotheses if players agree that one family is missing. This could be part of future developments of the game.

The integration of card-by-card feedback aligns with prior research emphasizing the value of formative assessment in CR education (Larsen et al., 2008). Feedback can relate to any aspect of the students' CR, such as their applied understanding of Hypothesis Families, their recognition of incomplete information, their hypotheses and their justification, their knowledge of the essential theoretical concepts inherent in the case study, etc. More broadly, studies have shown that interactive exchanges during CR practice allow previous knowledge to be reactivated, newly acquired information to be incorporated and integrated into increasingly structured and complex clinical scenarios (Chamberland et al., 2015; Gruppen, 2017; Thammasitboon et al., 2018). Therefore, immediate and focused feedback allows learners to identify cognitive gaps and refine their CR through iteration, fostering deeper understanding and encouraging reflective practice.

Immediate, targeted feedback allows learners to identify cognitive gaps and improve iteratively, promoting deeper understanding and reflective practice. This approach mirrors the structured feedback mechanisms in SCT but innovates by embedding it in an engaging SG format (Connolly et al., 2012). Using the same operating principles as physical games, SCALENEo effectively integrates these principles, encouraging learners' motivation and active participation. Indeed, the physical game showed that players globally felt motivated and encouraged (Hage et al., 2023), but in the context of this article, the role of player profiles (e.g., achievers vs. explorers), which is closely linked to gamification, was not taken into account. This is a potential avenue for future research (Vergara et al., 2023). To operationalize player profiles, we plan to administer established game-experience/flow instruments (e.g., those used in preliminary SCALENEo testing) and link them to anonymized usage telemetry (replay frequency, time on task, or error patterns) to investigate how engagement styles relate to CR performance.

4.2. Scoring CR

Our results demonstrate the value of scoring as a secondary, rather than primary, feedback mechanism. While scoring is not central in face-to-face sessions led by expert teachers, it is essential for unsupervised play in SCALENEo. Although we believe that face-to-face exchanges between students and experts will lead to a deeper understanding of CR than step-by-step online feedback and scoring, our proposal at least provides information that allows unsupervised players to self-assess their CR. Because the current algorithm scores are in concordance with information selection and categorization from expert references but cannot yet evaluate the quality or logic linking those choices, we plan to extend SCALENEo analytics to include coded chat exchanges and short think-aloud submissions to deepen interactional feedback when expert facilitation is unavailable.

We have shown that our scoring method for CR was sensitive to the percentage of errors made in the analysis of a clinical case: The more errors, the lower the score. This does not prove that our SG improves CR in real-life players, but it proves that a player making fewer errors will obtain a better score. So, our scoring method for CR may a priori be used to assess progress in CR, although normative values defining what constitutes a "good" CR still need to be established.

Global scoring reinforces cautious decision-making, rewarding those who analyze cases methodically rather than guessing. This approach aligns with educational theories that emphasize the importance of careful hypothesis testing (Larsen et al., 2008) and the mitigation of cognitive biases in clinical judgment (Croskerry, 2009b). It is particularly suited for intermediate learners who possess foundational knowledge of hypothesis families and are ready to refine their skills further. Because physiotherapy requires a broad base of skills and knowledge to enable the correct identification of clinical priorities and the application of appropriate medical interventions, we are satisfied with our scoring system that appears to favor caution: the risk-taking player always scores lower than the cautious player in this approach. This may teach students to always favor caution when assessing a clinical case, hoping that they can replicate this with real patients. The study of score evolution versus error rates underscores the importance of preliminary instruction in hypothesis families. The quick increase in scores for players with error rates below 50% suggests that SCALENEo is most effective when integrated into a broader curriculum with pre-game theoretical modules. This structured preparation ensures learners can engage meaningfully with the game's content, supporting both knowledge construction and skill application (Gorbanev et al., 2018). This structuring should increase players' motivation, particularly by encouraging them to replay games. This may be partly explained by two factors: the feeling of competence and the feeling of control over the task to be performed (Pelaccia & Viau, 2017; Viau, 2009). Although not directly measured in our study, these motivational dimensions are worth considering as potential mechanisms underlying student/learner engagement and should be explored in future research.

4.3. Limitations

Despite its strengths, certain limitations of our scoring method for the SCALENEo game must be acknowledged. We did not collect learner perception data (e.g., perceived openness of response options or frustration with closed item sets), which will be a priority for future comparative studies of SCALENEo and traditional SCT formats. The reliance on expert-defined reference games assumes homogeneity in CR, potentially overlooking variations in reasoning styles or regional practice differences. Future developments could explore adaptive feedback and dynamic difficulty adjustments to make the game more inclusive for novice learners while maintaining its value for intermediate and advanced users. Incorporating design principles from Koivisto et al. (2018) could further refine the game's accessibility and adaptability: as mentioned above, depending on the clinical case, different families of hypotheses could be deemed most important, and the score can, therefore, be adaptive. Additionally, the online platform offers significant opportunities for such research. Anonymized gameplay data could enable large-scale analyses of CR patterns across different demographics and educational backgrounds. Such insights could guide targeted instructional material development and address gaps in the existing curricula. For example, differences in reasoning approaches between novices and experts, as outlined by Croskerry (2009b), could inform personalized feedback mechanisms in SCALENEo.

4.4. Perspectives

SCALENEo's approach has broader implications beyond musculoskeletal physiotherapy. First, by adapting the SG to other domains, such as neurological or pediatric physiotherapy, its benefits could be extended to a wider range of learners. This scalability highlights its potential as a transformative tool in healthcare education. Furthermore, its emphasis on mitigating cognitive biases and encouraging cautious decision-making underscores its alignment with the principles of safe and effective patient care (Croskerry, 2009b). Second, it would be valuable to explore how learners perceive the SCALENEo experience in comparison to the traditional SCT. Specifically, examining their views on the degree of autonomy, cognitive engagement, and flexibility in hypothesis formulation could shed light on the educational impact of each format. Such insights may inform future iterations of the game and guide design choices that better balance structured support with open reasoning opportunities. Third, the future of providing in-depth, formative feedback on CR may lie more in the integration of artificial intelligence (AI) in complement to SG. AI holds significant potential to meaningfully personalize the learning process by adapting to individual learners' needs and performance. According to Tekin et al., AI demonstrates promise as a complementary tool for structured clinical evaluations, particularly in domains involving visually mediated clinical skills (Tekin et al., 2025). However, its reliability varies depending on the perceptual demands of the skill being assessed. The consistently higher and more uniform scores generated by AI systems indicate potential for standardization. Nonetheless, further refinement is necessary to ensure accurate assessment of skills that rely heavily on verbal communication and auditory cues. Although AI models are not yet equipped to replace human experts (Sallam & Abouzeid, n.d.), they can serve as valuable educational tools, supporting the development of reasoning skills, guiding learners through complex cases, and gradually narrowing the gap between students and expert clinicians.

5. Conclusions

SCALENEo scoring approach combines the strengths of SG with the rigor of established CR assessment methods. Its innovative approach to formative feedback, combined with an engaging and accessible online platform, positions it as a valuable tool for improving CR skills in physiotherapy education. By addressing cognitive biases, leveraging gamification principles, and integrating adaptive feedback, SCALENEo bridges the gap between knowledge and application in a way that traditional methods often fail to achieve. Our findings suggest that the adapted SCT scoring method implemented in the Happy Families game can differentiate between reasoning strategies, indicating its potential as a valid assessment tool for CR. While the current results do not directly demonstrate improved patient outcomes, they support the use of this tool to enhance CR training, which may ultimately contribute to better clinical decision-making in physiotherapy care.

We believe that our SG can be used at several stages during physiotherapy training. A first opportunity is when students complete their first physiotherapy internship, typically during their third year. At this point, they are considered to have the theoretical knowledge necessary to analyze a clinical case, and therefore to have a satisfactory understanding of the families of hypotheses. Our SG could then be used in preparation for the internship (to help structure the analysis of a patient's record), or after it, for example, by re-analyzing the case of a patient encountered during the internship. A second opportunity could be to use the SG at the beginning of the training, e.g., end of first year, as a motivational tool. After a quick definition of the families of hypotheses, students would be invited to analyze a simple clinical case and thus gradually understand the necessity of the courses that will be taught to them later.

Author Contributions: Conceptualization, R.H., F.D., S.D., W.E., S.L., J.-C.S., M.A.J. and F.B.; methodology, R.H., F.D., J.D.N., S.L. and F.B.; software, S.D. and W.E.; validation, R.H., F.D., S.D., W.E. and F.B.; formal analysis, F.D., J.D.N., S.D. and W.E.; investigation, F.D., W.E., S.L., J.-C.S. and F.B.; resources, R.H., J.D.N. and J.-C.S.; data curation, J.D.N., S.D. and W.E.; writing—original draft, R.H., F.D., W.E. and F.B.; writing—review and editing, R.H., F.D., J.D.N., S.L., J.-C.S., M.A.J. and F.B.; visualization, J.D.N. and M.A.J.; supervision, F.D., S.L., J.-C.S., M.A.J. and F.B.; project administration, F.D., S.L., J.-C.S. and F.B.; funding acquisition, F.D., S.L. and F.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by ERASMUS+, grant number 22PCS0007.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All generated data are available in Appendix A.

Acknowledgments: The authors thank Louis-David Beaulieu (Université du Québec à Chicoutimi-UQaC) for his feedback on the scoring algorithm and his involvement during test phases of the online application.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

AI	Artificial Intelligence
CR	Clinical Reasoning
SCALENEo	Smart ClinicAL rEasoning iN physiothErapy
SCT	Script Concordance Test
SG	Serious Game

Appendix A

Table A1. Numerical results obtained in the case where items were randomly removed. Average scores (of 20 generated games) are given with their standard deviation. The last two columns give the minimal and maximal scores that were generated.

Error (%)	Average Score	Score Standard Deviation	Min	Max
0	100	0	100	100
5	97.80	2.07	93	100
10	93.30	2.47	88	97
15	90.95	5.38	76	99
20	84.95	7.54	65	94
25	82.65	6.70	63	91
30	77.65	7.88	63	93
35	72.45	9.26	53	89
40	69.50	9.49	49	88
45	58.80	11.28	27	75
50	50.75	11.65	33	75
55	48.40	14.20	25	73
60	36.85	11.54	18	60
65	34.25	8.42	21	47
70	26.40	7.69	10	43
75	21.50	4.49	15	33
80	20	6.01	10	35
85	15.95	4.33	10	25
90	11	4.42	0	16
95	5.50	3.10	0	8
100	0	0	0	0

Table A2. Same as Table A1 in the case where items were randomly replaced.

Error (%)	Average Score	Score Standard Deviation	Min	Max
0	100	0	100	100
5	94.25	2.71	89	99
10	89.80	4.94	80	98
15	83.15	6.83	69	95
20	76.05	7.97	65	89
25	69.70	8.77	50	86
30	67.60	6.68	53	80

Table A2. *Cont.*

Error (%)	Average Score	Score Standard Deviation	Min	Max
35	60.45	9.83	36	76
40	54.75	8.43	38	67
45	45.00	9.88	26	62
50	39.30	10.42	23	63
55	35.45	8.99	21	53
60	30.35	10.20	13	53
65	28.15	11.88	12	62
70	20.60	10.42	6	45
75	15.35	8.59	0	38
80	9.10	5.25	0	23
85	11.60	7.35	0	25
90	3.35	6.11	0	24
95	2.85	5.21	0	20
100	0.40	1.19	0	5

References

Barlow, S. (2012). *Physiotherapy outpatient's chronic pain management . . . . . .Realizing the potential*. NSW Government.

Blanié, A., Amorim, M.-A., & Benhamou, D. (2020). Comparative value of a simulation by gaming and a traditional teaching method to improve clinical reasoning skills necessary to detect patient deterioration: A randomized study in nursing students. *BMC Medical Education*, *20*(1), 53. [CrossRef]

Bowen, J. L. (2006). Educational strategies to promote clinical diagnostic reasoning. *The New England Journal of Medicine*, *355*(21), 2217–2225. [CrossRef]

Brentnall, J., Thackray, D., & Judd, B. (2022). Evaluating the clinical reasoning of student health professionals in placement and simulation settings: A systematic review. *International Journal of Environmental Research and Public Health*, *19*(2), 936. [CrossRef]

Cant, R. P., & Cooper, S. J. (2010). Simulation-based learning in nurse education: Systematic review. *Journal of Advanced Nursing*, *66*(1), 3–15. [CrossRef]

Chamberland, M., Mamede, S., St-Onge, C., Setrakian, J., Bergeron, L., & Schmidt, H. (2015). Self-explanation in learning clinical reasoning: The added value of examples and prompts. *Medical Education*, *49*(2), 193–202. [CrossRef] [PubMed]

Charlin, B., Gagnon, R., Pelletier, J., Coletti, M., Abi-Rizk, G., Nasr, C., Sauvé, Év., & Van Der Vleuten, C. (2006). Assessment of clinical reasoning in the context of uncertainty: The effect of variability within the reference panel. *Medical Education*, *40*(9), 848–854. [CrossRef] [PubMed]

Charlin, B., Roy, L., Brailovsky, C., Goulet, F., & Van Der Vleuten, C. (2000). The script concordance test: A tool to assess the reflective clinician. *Teaching and Learning in Medicine*, *12*(4), 189–195. [CrossRef] [PubMed]

Clark, I., & Dumas, G. (2015). Toward a neural basis for peer-interaction: What makes peer-learning tick? *Frontiers in Psychology*, *6*. [CrossRef]

Cohen Aubart, F., Papo, T., Hertig, A., Renaud, M.-C., Steichen, O., Amoura, Z., Braun, M., Palombi, O., Duguet, A., & Roux, D. (2021). Are script concordance tests suitable for the assessment of undergraduate students? A multi-center comparative study. *La Revue de Médecine Interne*, *42*(4), 243–250. [CrossRef]

Connolly, T. M., Boyle, E. A., MacArthur, E., Hainey, T., & Boyle, J. M. (2012). A systematic literature review of empirical evidence on computer games and serious games. *Computers & Education*, *59*(2), 661–686. [CrossRef]

Croskerry, P. (2009a). A universal model of diagnostic reasoning. *Academic Medicine*, *84*(8), 1022–1028. [CrossRef]

Croskerry, P. (2009b). Clinical cognition and diagnostic error: Applications of a dual process model of reasoning. *Advances in Health Sciences Education*, *14*(S1), 27–35. [CrossRef]

Daniel, M., Rencic, J., Durning, S. J., Holmboe, E., Santen, S. A., Lang, V., Ratcliffe, T., Gordon, D., Heist, B., Lubarsky, S., Estrada, C. A., Ballard, T., Artino, A. R., Sergio Da Silva, A., Cleary, T., Stojan, J., & Gruppen, L. D. (2019). Clinical reasoning assessment methods: A scoping review and practical guidance. *Academic Medicine*, *94*(6), 902–912. [CrossRef]

Deschênes, M.-F., Charlin, B., Gagnon, R., & Goudreau, J. (2011). Use of a script concordance test to assess development of clinical reasoning in nursing students. *The Journal of Nursing Education*, *50*(7), 381–387. [CrossRef]

Dory, V., Gagnon, R., Charlin, B., Vanpee, D., Leconte, S., Duyver, C., Young, M., & Loye, N. (2016). In brief: Validity of case summaries in written examinations of clinical reasoning. *Teaching and Learning in Medicine*, *28*(4), 375–384. [CrossRef] [PubMed]

Dory, V., Gagnon, R., Vanpee, D., & Charlin, B. (2012). How to construct and implement script concordance tests: Insights from a systematic review. *Medical Education*, *46*(6), 552–563. [CrossRef] [PubMed]

Edwards, I., Jones, M., Carr, J., Braunack-Mayer, A., & Jensen, G. M. (2004). Clinical reasoning strategies in physical therapy. *Physical Therapy*, *84*(4), 312–330. [CrossRef] [PubMed]

Finucane, L. M., Downie, A., Mercer, C., Greenhalgh, S. M., Boissonnault, W. G., Pool-Goudzwaard, A. L., Beneciuk, J. M., Leech, R. L., & Selfe, J. (2020). International framework for red flags for potential serious spinal pathologies. *Journal of Orthopaedic & Sports Physical Therapy*, *50*(7), 350–372. [CrossRef]

García-Redondo, P., García, T., Areces, D., Núñez, J. C., & Rodríguez, C. (2019). Serious games and their effect improving attention in students with learning disabilities. *International Journal of Environmental Research and Public Health*, *16*(14), 2480. [CrossRef]

Gorbanev, I., Agudelo-Londoño, S., González, R. A., Cortes, A., Pomares, A., Delgadillo, V., Yepes, F. J., & Muñoz, Ó. (2018). A systematic review of serious games in medical education: Quality of evidence and pedagogical strategy. *Medical Education Online*, *23*(1), 1438718. [CrossRef]

Gruppen, L. D. (2017). Clinical reasoning: Defining it, teaching it, assessing it, studying it. *Western Journal of Emergency Medicine*, *18*(1), 4–7. [CrossRef]

Hage, R., Fourré, A., Ramonfosse, L., Leteneur, S., Jones, M., & Dierick, F. (2023). Description and rules of a new card game to learn clinical reasoning in musculoskeletal physiotherapy. *Journal of Manual & Manipulative Therapy*, *31*(4), 287–296. [CrossRef]

Haoran, G., Bazakidi, E., & Zary, N. (2019). Serious games in health professions education: Review of trends and learning efficacy. *Yearbook of Medical Informatics*, *28*(01), 240–248. [CrossRef] [PubMed]

Hickey, D. T. (1997). Motivation and contemporary socio-constructivist instructional perspectives. *Educational Psychologist*, *32*(3), 175–193. [CrossRef]

Huhn, K., Gilliland, S. J., Black, L. L., Wainwright, S. F., & Christensen, N. (2019). Clinical reasoning in physical therapy: A concept analysis. *Physical Therapy*, *99*(4), 440–456. [CrossRef]

Jones, M. A. (1992). Clinical reasoning in manual therapy. *Physical Therapy*, *72*(12), 875–884. [CrossRef]

Jones, M. A., & Rivett, D. A. (2004). *Clinical reasoning for manual therapists*. (Butterworth Heinemann). Elsevier Science Limited.

Jones, M. A., Rivett, D. A., & Moore, A. (2019). *Clinical reasoning in musculoskeletal practice* (2nd ed.). Elsevier.

Koivisto, J.-M., Haavisto, E., Niemi, H., Haho, P., Nylund, S., & Multisilta, J. (2018). Design principles for simulation games for learning clinical reasoning: A design-based research approach. *Nurse Education Today*, *60*, 114–120. [CrossRef]

Kojich, L., Miller, S. A., Axman, K., Eacret, T., Koontz, J. A., & Smith, C. (2024). Evaluating clinical reasoning in first year DPT students using a script concordance test. *BMC Medical Education*, *24*(1), 329. [CrossRef]

Kononowicz, A. A., Woodham, L. A., Edelbring, S., Stathakarou, N., Davies, D., Saxena, N., Tudor Car, L., Carlstedt-Duke, J., Car, J., & Zary, N. (2019). Virtual patient simulations in health professions education: Systematic review and meta-analysis by the digital health education collaboration. *Journal of Medical Internet Research*, *21*(7), e14676. [CrossRef]

Krishnamurthy, K., Selvaraj, N., Gupta, P., Cyriac, B., Dhurairaj, P., Abdullah, A., Krishnapillai, A., Lugova, H., Haque, M., Xie, S., & Ang, E.-T. (2022). Benefits of gamification in medical education. *Clinical Anatomy*, *35*(6), 795–807. [CrossRef]

Lapkin, S., Levett-Jones, T., Bellchambers, H., & Fernandez, R. (2010). Effectiveness of patient simulation manikins in teaching clinical reasoning skills to undergraduate nursing students: A systematic review. *Clinical Simulation in Nursing*, *6*(6), e207–e222. [CrossRef]

Larsen, D. P., Butler, A. C., & Roediger, H. L., Iii. (2008). Test-enhanced learning in medical education. *Medical Education*, *42*(10), 959–966. [CrossRef]

Lubarsky, S., Charlin, B., Cook, D. A., Chalk, C., & van der Vleuten, C. P. M. (2011). Script concordance testing: A review of published validity evidence. *Medical Education*, *45*(4), 329–338. [CrossRef] [PubMed]

Min Simpkins, A. A., Koch, B., Spear-Ellinwood, K., & St. John, P. (2019). A developmental assessment of clinical reasoning in preclinical medical education. *Medical Education Online*, *24*(1), 1591257. [CrossRef] [PubMed]

O'Brien, S. R., Dillon, N., Linsky, M., Lagueras, K., Uhl, J., Conroy, S., Kiley, A., Lanzafame, N., States, E., & Sturm, E. (2023). Initial validation of a script concordance test to measure the development of clinical reasoning among physical therapy residents. *The Journal of Clinical Education in Physical Therapy*, *5*, 9014. [CrossRef]

Pelaccia, T., & Viau, R. (2017). Motivation in medical education*. *Medical Teacher*, *39*(2), 136–140. [CrossRef]

Pinnock, R., Young, L., Spence, F., Henning, M., & Hazell, W. (2015). Can think aloud be used to teach and assess clinical reasoning in graduate medical education? *Journal of Graduate Medical Education*, *7*(3), 334–337. [CrossRef]

Qiao, J., Huang, C.-R., Liu, Q., Li, S.-Y., Xu, J., & Ouyang, Y.-Q. (2023). Effectiveness of nonimmersive virtual reality simulation in learning knowledge and skills for nursing students: Metaanalysis. *Clinical Simulation in Nursing*, *76*, 26–38. [CrossRef]

Roediger, H. L., & Karpicke, J. D. (2006). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science*, *17*(3), 249–255. [CrossRef]

Rushton, D. A., Beeton, D. K., Jordaan, D. R., Langendoen, M. J., Levesque, M. L., & Maffey, M. L. (2016). *Part A: Educational standards 2016*. International Federetion for Orthopeadics Manipulative Therapists.

Sallam, M. A., & Abouzeid, E. n.d. Introducing AI as members of script concordance test expert reference panel: A comparative analysis. *Medical Teacher*, 1–8. [CrossRef]

Schuwirth, L. (2009). Is assessment of clinical reasoning still the Holy Grail? *Medical Education*, *43*(4), 298–300. [CrossRef] [PubMed]

Smart, K., & Doody, C. (2007). The clinical reasoning of pain by experienced musculoskeletal physiotherapists. *Manual Therapy*, *12*(1), 40–49. [CrossRef] [PubMed]

Tekin, M., Yurdal, M. O., Toraman, Ç., Korkmaz, G., & Uysal, İ. (2025). Is AI the future of evaluation in medical education? AI vs. Human evaluation in objective structured clinical examination. *BMC Medical Education*, *25*(1), 641. [CrossRef] [PubMed]

Thammasitboon, T., Jj, R., Rl, T., Apj, O., M, S., & G, D. (2018). The assessment of reasoning tool (ART): Structuring the conversation between teachers and learners. *Diagnosis (Berlin, Germany)*, *5*(4), 197–203. [CrossRef]

Tubelo, R. A., Portella, F. F., Gelain, M. A., De Oliveira, M. M. C., De Oliveira, A. E. F., Dahmer, A., & Pinto, M. E. B. (2019). Serious game is an effective learning method for primary health care education of medical students: A randomized controlled trial. *International Journal of Medical Informatics*, *130*, 103944. [CrossRef]

Vergara, D., Antón-Sancho, Á., & Fernández-Arias, P. (2023). Player profiles for game-based applications in engineering education. *Computer Applications in Engineering Education*, *31*(1), 154–175. [CrossRef]

Viau, R. (2009). *La motivation en contexte scolaire* (5e éd.). De Boeck.

Young, M., Thomas, A., Gordon, D., Gruppen, L., Lubarsky, S., Rencic, J., Ballard, T., Holmboe, E., Da Silva, A., Ratcliffe, T., Schuwirth, L., & Durning, S. J. (2019). The terminology of clinical reasoning in health professions education: Implications and considerations. *Medical Teacher*, *41*(11), 1277–1284. [CrossRef]