

Development of Situational Judgment Tests in Interprofessional Health Education

Fernanda Drummond Ruas Gaspar, Gardênia da Silva Abbad

Universidade de Brasília, Graduate Programme in Social Psychology of Work and Organisations, Brasília, Distrito Federal, Brazil
Email: fernandagaspar1202@gmail.com

How to cite this paper: Gaspar, F. D. R., & da Silva Abbad, G. (2024). Development of Situational Judgment Tests in Interprofessional Health Education. *Creative Education*, 15, 21-45.

<https://doi.org/10.4236/ce.2024.151002>

Received: December 12, 2023

Accepted: January 22, 2024

Published: January 25, 2024

Copyright © 2024 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

Situational Judgment Tests (SJTs) have been considered a valuable strategy to assess attitudinal skills learning in practical scenarios. The objectives of this study are: 1) to describe the process of development and semantic and content evaluation of a SJT in the work of health care residents; 2) to describe the development of a correction sheet for the STJ (answer sheet); and 3) to test the equivalence of complexity of the content of the SJT items. The data from 15 interviews conducted with preceptors and both medical and non-medical residents were analyzed by a group of 6 researchers. They identified gaps in the predominant social skills necessary for collaborative work and training of health residents: understanding roles and responsibilities, debate of different opinions, and collaborative decision-making. Based on these three skills, six real-life situations in multidisciplinary health settings and four open-ended questions were developed for residents to solve. The test content underwent both semantic and content validation. Three tests were selected for a pilot application. The scores generated from answers corrected by two researchers were submitted to statistical analyses to assess content complexity equivalence. Participants underperformed in the SJT, suggesting learning gaps in the assessed social skills. There were no significant differences in scoring, confirming the content complexity equivalence among different situations. The administration of these SJTs has the potential to enhance resident education and interprofessional health training, fostering socio-affective interactions within collaborative work contexts.

Keywords

Situational Judgment Test, Case-Based Learning, Social Skill, Health Residency

1. Introduction

Recently, Situational Judgment Tests (SJTs) have gained visibility in the field of

learning assessment due to their practical, accurate, and sensitive approach to the specificities of different educational and professional contexts (Barron et al., 2022; Cualheta & Abbad, 2021; Weng et al., 2018; Wolcott et al., 2020). This type of test requires respondents to describe how they would act in a particular situation or dilemma (Cualheta & Abbad, 2021). In the healthcare domain, these tests are viewed as a cost-effective and appropriate assessment strategy, encompassing both aspects of clinical reasoning and behavioral elements of interactions with patients, families, and professionals (Kiessling et al., 2016). Furthermore, SJTs prove more effective in measuring personal and intrapersonal skills, as they mitigate bias and validity issues, reducing the likelihood of misrepresentation, and enabling the assessment of changes in individuals' behaviours and attitudes (Anderson et al., 2017).

Employing active methodologies in training (Cox et al., 2017) and learning assessment tools (Wolcott et al., 2020), including practical and realistic formats, such as SJTs, may enhance the positive effects of interprofessional education endeavors. The training and practical scenario of health residency seem to be suitable for scientific investigations on the construction, application, and search for evidence of validity of these tests.

Residents are constantly immersed in the practice setting, enabling significant learning and immediate application of both technical and non-technical skills essential for comprehensive and high-quality healthcare. Non-technical or social skills encompass cognitive and, primarily interpersonal aspects (Del Prette & Del Prette, 2018) that complement the clinical skills necessary for ensuring patient safety (Chang et al., 2019; Collins et al., 2021; Mata et al., 2021; Reeves et al., 2013). Communication, for instance, stands as one of these non-technical skills, carrying a heightened risk of errors and failures in the hospital environment (Riley et al., 2011), thereby reinforcing the need for investment in continuous and effective training, development and education actions throughout the careers of health professionals (Jin et al., 2022; Lamba et al., 2016).

Few studies were identified in the literature on learning assessment strategies attuned to the impact of interventions aimed at enhancing affective skills in the work context of health care residents, such as those mentioned by Thistlethwaite et al. (Thistlethwaite et al., 2010). These strategies encompass elements such as teamwork, role definition, communication, feedback, learning, reflective thinking, patient-centred focus, and ethical conduct. Some authors deem it essential to undertake research focused on the development and evaluation of learning assessment instruments tailored for use within collaborative settings comprising varied health care professionals (Kang et al., 2022; Thistlethwaite et al., 2010).

The development of SJTs aimed at assessing social skills acquisition is substantiated by the relevance of these skills in the practice and training of the health residents, along with the gaps identified in the literature regarding learning assessment instruments validated for their psychometric validity and reliability. Consequently, this study holds three primary objectives: 1) to describe the con-

struction process and assess the semantic and content aspects of SJTs within the residents' work context; 2) to create an answer sheet for SJTs (template); 3) to assess the equivalence of content complexity across SJTs items.

2. Learning Assessment Instruments in the Health Context

Developing a learning assessment instrument demands adherence to multiple quality criteria to ensure valid, reliable, and comprehensive measurement of knowledge, skills, and attitudes exhibited in simulated or real situations (Lineberry et al., 2013). Health education curricula require improved assessments with evidence of validity. While there are attitudinal assessments for teamwork, only a few involve direct observation of clinical behaviours, such as the Teamwork Skills (CATS), which assesses communication competence, and the Teamwork Mini-Clinical Evaluation Exercise (T-MEX), which develops active participation focused on patient care. Assessments should also be ongoing processes, rather than taking place at a single moment, integrating debriefing and feedback stages throughout the training actions (Havyer et al., 2016).

The Readiness for Interprofessional Learning Scale (RIPLS) is an example of a tool to assess specific attitudes related to collaborative learning between different health subjects and professions. However, The RIPLS has been criticized for its subjective character and its limited correlation with the actual performance of health teams. Concerns have also arisen regarding its factorial structure and imprecision in its measurement focus. As a result, the scale has not been recommended for the evaluation of interventions and to compare different training results (Mahler et al., 2015). Nevertheless, Peduzzi et al. (Peduzzi et al., 2015) argue that the Portuguese-translated version of the scale may be useful for policymaking, planning interprofessional education programs, and understanding the interprofessional learning experiences of health students.

The review conducted by Kang et al. (Kang et al., 2022) presented seven instruments designed for the assessment of collaborative efforts within health care teams. Most of these instruments consist of objective and quantitative items related to attributes that lead to the socio-affective interactions established between health care professionals, including communication, role definition, and decision-making.

The Objective Structured Clinical Examination (OSCE) is regarded as an effective assessment method in health education. Learners are immersed in clinical scenarios that closely resemble real work situations, and both their technical and non-technical skills are assessed through direct observation, supported by the use of checklists (Lamba et al., 2016).

In health education, where there is a multitude of assessment instruments and strategies, it is crucial to assess the advantages and disadvantages of each method. The choice of data collection strategies should align with the study's focus on quantitative, qualitative, or a combination of data and the specific dimension

of the phenomenon under examination (Shrader et al., 2017). Regardless of the wide variety of scientific instruments available, a careful evaluation of their adequacy and usefulness is recommended, since in some cases it may be more appropriate to build specific instruments that are tailored to the specific research context (Cualheta & Abbad, 2021).

There are also other aspects that must be considered during the process of selection, construction and application of a research instrument, such as the presence of items that express observable teamwork behaviours, taking into account the current context of its application, and providing evidence of validity. Multiple assessment sources, such as self-report measures and evaluations conducted by observers or evaluators, should also be used to reduce bias and subjectivity in the ratings, align the instrument items with the target audience's characteristics, and train the raters and observers who will administer the instrument and evaluate participants' responses (Kang et al., 2022; Marlow et al., 2018).

Instruments based on real working situations are more appropriate for assessing (cognitive and social) skill learning than traditional declarative knowledge instruments, which focus on content recall and not the demonstration of behavioral change. The shape of these traditional instruments does not stimulate the analytical capacity of the apprentice on critical and recipient situations of the workplace, contributing to the maintenance of relevant professional competence gaps.

Situational Learning Testing

While several instruments for measuring clinical and technical skills are commonly used in healthcare training, educators and instructional planners continue to face challenges when it comes to assessing and measuring non-technical skills, behaviours, and attitudes. The SJT emerged as a valuable methodological approach for assessing individuals' underlying skills and attributes when faced with realistic work scenarios, depicting dilemmas or problems that demand the application of knowledge, skills, and attitudes (Wolcott et al., 2020). Test items may be presented in written, verbal or video formats and may contain multiple appropriate responses (Christian et al., 2010; Reed et al., 2022; Patterson et al., 2016; Smith et al., 2020). These scenarios are meticulously crafted based on a thorough analysis of the context, function, or activity under assessment and are ideally developed in collaboration with experts in the subject matter to ensure accuracy and alignment with real assessment requirements (Cechella et al., 2021; Cualheta & Abbad, 2021; Patterson et al., 2016).

Reed et al. (Reed et al., 2022) conducted a literature review, revealing a positive correlation between the performance of professionals in situational tests and the future performance of health care professionals at work. This finding can be valuable in informing educational strategies aimed at cultivating attitudes and enhancing social and affective skills. Notably, social skills present a challenge for objective measurement, making them a particularly suitable domain for the ap-

plication of SJTs. These tests can measure implicit prosocial traits, as well as individuals' beliefs and values regarding behaviors and situations. They contribute to teaching the expression of attitudes across different social settings, ranging from pleasant expressions, such as helping other people, to unpleasant expressions, such as prioritizing one's interests over those of others (Smith et al., 2020).

According to Lievens & Sackett's (Lievens & Sackett, 2012) research, the application of SJTs using videos could be associated with both academic and post-academic success criteria. Their study revealed that some physicians, despite having little technical experience, but with procedural knowledge related to effective behaviours in interpersonal relationships, achieved high scores in admission tests. According to this study, SJT scores serve as important predictors of the actual behaviour exhibited by these physicians in future interpersonal situations. One notable advantage of this type of test is its ability to assess complex constructs among large groups of examinees, making it an appealing alternative or complement to more resource-intensive assessment methods such as situational interviews or objective structured clinical examinations (Lievens & Sackett, 2012).

The process of constructing high-quality learning SJTs involves several key steps, including 1) developing test specifications, 2) designing scenarios and response options, 3) establishing key answers and scoring methods, 4) building the test, 5) conducting pilot testing, 6) performing psychometric analyses, and 7) maintaining and regularly updating the question bank. It is recommended to engage experts in the relevant subject area and perform a thorough analysis of the work environment to identify the skills, tasks, and other characteristics that should be included in the assessment (Cechella et al., 2021; Cualheta & Abbad, 2021; Reed et al., 2022). The use of critical incidents is a valuable technique for capturing realistic work situations. Ideally, the tests should be concise and objective, considering the time constraints often faced by healthcare professionals when completing these instruments. However, when reducing the number of test questions one should consider that the validity of the instrument should be preserved (Kang et al., 2022).

The SJT also presents some disadvantages, including its complex construction process, and the need for clinical educators to be well-versed in aspects like validity and reliability assessments (Reed et al., 2022). Additionally, defining the constructs for assessing non-technical skills through SJT can be challenging and requires careful scrutiny to ensure accuracy and validity (Reed et al., 2022; Wolcott et al., 2020).

In summary, SJT: 1) is an important learning tool in student and professional training for the labour market, 2) is based on real situations and observable behaviours, 3) stimulates reflections and analysis on dilemmas extracted from circumstances of daily work, 4) can evaluate knowledge, skills and attitudes and 5) follow a rigorous process of elaboration, with the support of experts and diagnoses that reflect reality.

Faced with the challenges reported in the design and implementation of SJT, this study intends to provide a comprehensive account of 1) the development process of SJT and its semantic and content evaluation, 2) the creation of the template for test correction and, 3) the assessment of equivalence in the complexity of SJT item content.

3. Method

This study employs a mixed sequential exploratory approach, associating both qualitative and quantitative data and multiple sources of data analysis. This approach enhances our understanding of the research problem and contributes to the development of more contextually relevant instruments with methodological rigour (Creswell & Creswell, 2021; Levitt et al., 2018).

A total of 6 SJTs were developed to measure the acquisition of social skills among residents in the multi-professional health care context. These SJTs were chosen as the assessment instrument for evaluating social skills training among health care residents in Brazil. The development process, semantic, and content evaluation of these SJTs, and the creation of test correction templates comprised nine steps, which are described below (Figure 1), outlining participants, data collection, and analysis procedures.

The first stage comprised semi-structured interviews with 15 health professionals (4 residents and 11 preceptors of medical and multiprofessional residencies) at a Brazilian University Hospital. These participants were selected by convenience sampling. The aim was to collect critical incidents related to socio-affective skills employed by residents in their multi-professional work routine. In addition to recounting these critical incidents, respondents were asked to identify which social skills residents should enhance.

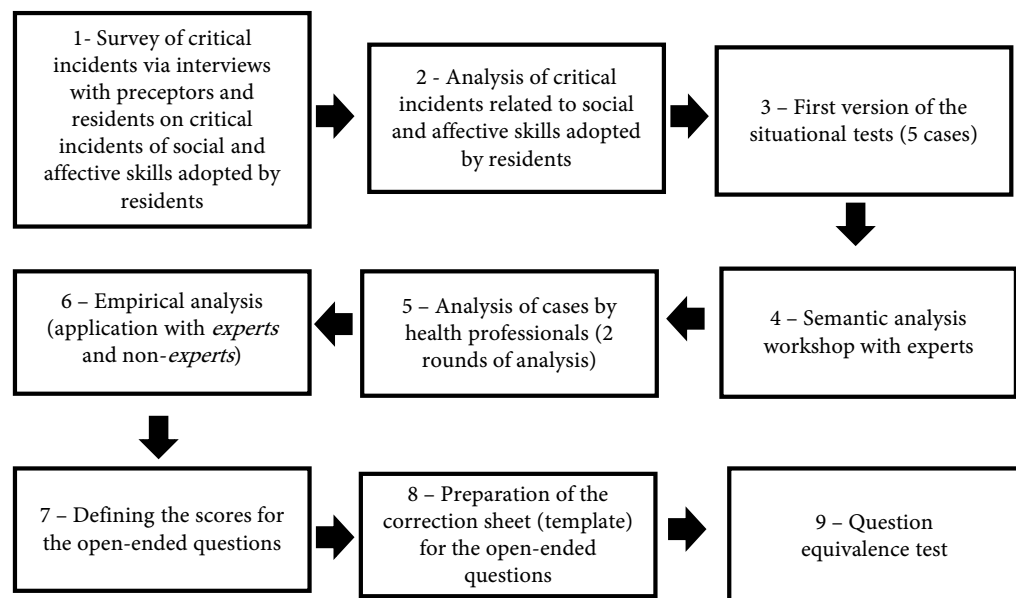


Figure 1. Stages of Construction and Search for Evidence of validity of the SJT.

In the second stage, the responses from the interviewees were transcribed verbatim and subjected to analysis by a team of six researchers. Initially, each researcher independently reviewed the transcriptions, identifying prominent social skills gaps essential for the collaborative work and training of health residents. To guide this process, researchers were provided with a list of social skills categories extracted from existing scientific literature (Del Prette & Del Prette, 2018; Kang et al., 2022). Subsequently, the research team reached a consensus, identifying the primary social skills learning gaps related to 1) understanding the roles and responsibilities of health professionals from different academic backgrounds, 2) considering different opinions, and 3) engaging in collaborative decision-making regarding patient care.

Subsequently, in the third stage of the study, a preliminary version of the tests was prepared based on the descriptions of situations that indicate gaps in social skills. This initial version consisted of six different cases, each depicting a specific problem situation that demanded the use of the previously identified three social skills gaps among residents. Additionally, each case included four open-ended questions as follows: 1) Identify the professions of the team members involved in patient X's care; 2) Describe three actions or behaviors presented by the team members that contributed to the issue faced by patient X; 3) Explain how the team members could have acted differently to avoid situations similar to that of patient X; and 4) Given the solutions indicated, describe how you would communicate with the team members to prevent future occurrences resembling that of patient X. Provide details on what you would say to the team and what the dialogue would look like.

The open-ended questions were identical in all cases and required the respondent to use a range of skills from simpler tasks, such as describing the professionals involved in the case, to more complex abilities, such as devising solutions to the case-related problems. The questions were formulated based on the learning gaps identified in the first stage of test construction and the instructional objectives of a training program on social skills for residents in the multi-professional context. To ensure uniformity, the complexity level of each question was based on Bloom's Taxonomy of Learning Outcomes (1956) while maintaining equivalence across the situational scenarios and the corresponding questions to be answered by the research participants.

After drafting the first version of the SJTs, a workshop was conducted online as part of the fourth stage to seek evidence of semantic validity. This workshop involved nine members of a research group with expertise in instrument construction and evaluation of educational programmes, representing diverse academic backgrounds (psychology, pedagogy, administration) and educational levels (complete graduation, incomplete graduation, complete undergraduate education, and undergraduate education in progress). Following initial instructions, participants were divided into three sub-groups, each consisting of three participants. Each subgroup was tasked with reviewing two cases and collaboratively answering five questions:

1) Is the case generally described with clear and accessible language? Suggest improvements if needed.

2) Are the four questions at the end of the case clear and accessible? Suggest improvements if needed.

3) Are there any grammatical or linguistic inconsistencies or errors in the Portuguese language? Please specify (paragraph and line).

4) Does the case lack any essential information for comprehension? Please specify.

5) Can any non-essential information be removed without affecting comprehension?

Following the discussions and questions completion, each subgroup presented their answers and improvement suggestions. These mainly involved correcting spelling errors and adding character information to enhance case comprehension. Additionally, the group suggested modifications to the four open-ended questions in each case. Questions “a” and “b” were adjusted to focus more on descriptive aspects and scenario mapping, while questions “c” and “d” were revised to involve higher complexity, emphasizing analysis and problem-solving skills.

After the experts’ analysis, the fifth stage involved seeking evidence of content validity. A total of 17 health professionals from different backgrounds (4 physicians, 3 nurses, 2 psychologists, 2 physiotherapists, 2 pharmacists, 1 dentist, 1 speech therapist, 1 nutritionist, 1 nursing technician) participated in this stage. Professionals were selected according to the researcher’s availability, and the criteria for participation included having experience working in a hospital setting (public or private) within multiprofessional health teams. Cases were distributed so that each professional analyzed content from at least two different situations. Their specific health training was considered during case allocation to ensure that the case would be analysed by professionals with different trainings and expertise. All professionals were instructed to read the cases and answer three questions: 1) Do you believe the cases have practical relevance? 2) Are there any theoretical or technical inconsistencies in the case descriptions? 3) Is there any missing information that would enhance understanding of the cases?

The answers to these questions were transcribed verbatim, organized in an Excel spreadsheet and analysed by a researcher. All 17 professionals indicated that all cases had practical pertinence. However, some professionals pointed out technical inconsistencies, such as the inappropriate use of technical terms, like “broncho-aspiration”, which was changed to “airway aspiration”. Some professionals also suggested providing more details about the patient’s condition.

After making adjustments to the cases, we conducted another round of content analysis with three health professionals to confirm that the changes did not compromise the technical description of the case. During this analysis, two pharmacists and a nursing technician disagreed on the technical consistency of case 5, which involved a situation where a pharmacist refused a medication pre-

scription from a physician. Since there was no consensus between the two pharmacists regarding the technical consistency of case 5, we decided to exclude this case from the study.

After completing the semantic and content analysis, the sixth stage involved conducting a pilot test with a group of health professionals. The main objective was to assess the consistency in the complexity of the test content. Three of the five cases presented in Appendix A were selected randomly: Case 2 (Failure to aspirate a patient's airways), Case 4 (Patient is discharged without Occupational Therapy guidelines), and Case 6 (Modification of a patient's diet). This pilot study also aimed to gather feedback from participants regarding item instructions and structure, in line with recommendations (Cualheta & Abbad, 2021).

The three test versions were applied in a single session with the health professionals in this sequence: T1 = Case 2; T2 = Case 4; and T3 = Case 6. The tests were transposed into a single online file using Google Forms and distributed to 37 professionals from 8 different health backgrounds, including experts and non-experts. While the target audience for the SJT is primarily residents, both medical and non-medical, the pilot application involved health professionals with varying levels of experience. This approach was chosen to collect a wide range of responses, both desirable and undesirable, to aid in the definition of scoring criteria (Cualheta & Abbad, 2021).

The professionals also selected by convenience were instructed to read the three cases described in the SJT and respond to the four open questions. Experts were identified as those professionals with a minimum of four years of experience in hospital multidisciplinary teams. All participants in the pilot test (Table 1) signed an Informed Consent Form (ICF) to ensure the confidentiality of shared information. The Research Ethics Committee at our institution (Faculdade de Saúde-Universidade de Brasília) approved the research (opinion number: 5.362.942)

Most participants were women (83.8%), held a medical degree (32.4%) and had up to 5 years of work experience (35.1%). Additionally, a significant proportion of participants met the criteria for classification as experts, with a minimum of 4 years of experience (70.2%).

The application of the situational tests allowed us to gather responses from health professionals, which facilitated the seventh stage—defining scoring criteria for each type of response to the four questions. To ensure that the test items were properly aligned with the complexity of the residents' learning needs, the behaviours described in the cases were analyzed according to Bloom's taxonomy (1956), which covers three dimensions of learning (cognitive, affective and psychomotor) and different levels of complexity. Expectations of correct answers for each item are presented in Appendix B.

All three tests were developed based on the knowledge, skills, and attitudes expected from the student (resident). Furthermore, the level of complexity was considered when formulating each question. These questions are open-ended, encouraging the resident to provide a comprehensive written response outlining

Table 1. Distribution of health professionals who responded the SJT pilot application.

Variable	Category	Frequency	Percentage
Gender	Women	31	83.8%
	Men	6	16.2%
Expertise	Expert	26	70.2%
	Non-Expert	10	29.8%
Academic Background	Medicine	12	32.4%
	Nursing	8	21.6%
	Psychology	5	13.5%
	Physiotherapy	3	8.1%
	Odontology	3	8.1%
	Nursing Technician	3	8.1%
	Pharmacy	2	5.4%
	Occupational Therapy	1	2.7%
Length of Experience	0 - 5 years	13	35.1%
	6 - 10 years	11	29.7%
	11 - 15 years	4	10.8%
	16 - 20 years	4	10.8%
	21 - 25 years	3	8.1%
	26 - 30 years	2	5.4%

Source: Prepared by the authors, N = 37.

their actions in the situations presented. In particular, question/item “a” holds a weight of 1.0 and requires the resident to identify the professions involved in the given case scenario. This question has a low level of complexity (level of knowledge – cognitive dimension, according to Bloom’s taxonomy, 1956). Its main objective is to verify whether the resident can identify the various professions within the multidisciplinary team that engage in daily socio-affective interactions in the context of collaborative work.

Question “b” has a weight of 2.0 and asks the resident to describe three behaviors presented by the team of health care professionals that contributed to the issue described in the case. This question has a higher degree of complexity compared to the previous one, as it requires the resident to analyze the case scenario (level of understanding—cognitive dimension and level of appreciation— affective dimension, according to Bloom’s taxonomy, 1956), focusing on distinguishing the tasks, responsibilities, and roles of each member of the multidisciplinary team, while concurrently supporting collaborative efforts.

Regarding question “c,” with a weight of 3.0, it prompts the residents to explain how the healthcare team professionals could have taken actions to prevent

situations such as those mentioned in the case (level of appreciation— affective dimension and level of evaluation—cognitive dimension, according to Bloom’s taxonomy, 1956). The item aims to assess shortcomings in socio-affective interactions within the workplace and the effects on patient care.

Finally, question “d”, with a weight of 4.0, instructs the resident to provide a step-by-step description of how they would communicate with their team members to address the problem outlined in the case. This question demands a high level of detail from the respondent, including information on “what” and “how” the dialogue with the team would unfold, assessing both the organization aspect (affective dimension) and the evaluative aspect (cognitive dimension), according to Bloom’s taxonomy (1956). The item aims to assess the resident’s ability to propose solutions, grounded in discussion with team members, and collaboratively make decisions on aspects regarding patient care.

Stage eight involved preparing the answer sheets (templates) based on the established criteria. An illustration of one such template (Question A, Case 2) is presented in Appendix C. Two researchers independently analyzed the test answers and assigned scores for each question according to the answer sheets (templates).

Subsequently, the researchers engaged in discussions to establish a consensus on the scores assigned in the pilot test. All the answers were discussed by both researchers to determine the ideal answers and establish consistent scoring criteria for each item. Given that the questions are open-ended and the skill content is extensive and dynamic, a wide range of correct answers can be expected. Therefore, the answer sheet (template) serves to reduce subjectivity and potential biases in the correction, as it incorporates well-defined scoring criteria.

In the ninth and last stage, an analysis was conducted to assess the equivalence of the complexity of the content in the items and questions of the three SJTs (cases 2, 4 and 6) administered in the pilot test and corrected based on the established criteria. The final scores of the participants in each of the tests were entered into the statistical software Jamovi (version 2.2.5). It is important to note that the three tests featured different problem situations but shared equivalent evaluation objectives and assessed skills. All three versions of the test focused on the same three social skills mentioned above: 1) understanding the roles and responsibilities of health care professionals from various academic backgrounds, 2) engaging in debates with different opinions, and 3) collaborative decision-making in patient care.

In addition to providing descriptive statistics such as mean, standard deviation, and minimum and maximum values, an evaluation was conducted by comparing the total scores assigned by evaluators to the professionals participating in the study across the three tests. Additionally, participants’ scores for each question were also compared. The Wilcoxon test was chosen for both comparisons due to the non-normal distribution of the data. This approach aimed to assess whether the contents of the SJTs were equivalent in complexity. Statistical significance was established at $p < 0.05$ (Tabachnick & Fidell, 2013).

4. Results

The three SJTs (T1, T2 and T3), corresponding to cases 2, 4 and 6, can be administered at various points in time, both before and after training, serving as assessment tools for evaluating the acquisition of social skills in medical and non-medical residents. Detailed descriptions of the three cases forming the SJTs (T1, T2 and T3) can be found in Appendix A.

The three SJTs describe problems faced by health care professionals in different contexts, including the ICU, outpatient clinic, and inpatient unit). These scenarios present socio-affective interaction problems involving at least two professionals from distinct academic backgrounds. To arrive at the correct solutions to these problems, professionals are expected to demonstrate their proficiency in social and affective skills derived from the assessment of learning needs. These skills encompass understanding roles and responsibilities, engaging in debates with different opinions, and making collaborative decisions. **Table 2** shows the total scores for the answers in the three SJTs (T1, T2 and T3), as well as individual scores for open-ended questions (“a”, “b”, “c” and “d”). It is important to note that participants could attain a maximum of 10 points in each test, and 1.0, 2.0, 3.0 and 4.0 in questions “a”, “b”, “c” and “d”, respectively.

Table 2. Mean scores by question.

	Mean (M)	Standard Deviation	Minimum	Maximum
Situational Judgment Test 1 (questions a, b, c, d)				
T1_a	0.68	0.25	0.00	1.00
T1_b	0.89	0.49	0.00	1.82
T1_c	1.67	0.79	0.00	3.00
T1_d	1.96	0.88	0.00	4.00
TOTAL	5.20	1.44	3.06	9.48
Situational Judgment Test 2 (questions a, b, c, d)				
T2_a	0.89	0.15	0.50	1.00
T2_b	0.97	0.38	0.00	1.71
T2_c	1.16	0.70	0.00	3.00
T2_d	1.82	0.89	0.00	3.50
TOTAL	4.84	1.47	1.41	8.06
Situational Judgment Test 3 (questions a, b, c, d)				
T3_a	0.86	0.20	0.00	1.00
T3_b	0.82	0.30	0.00	2.00
T3_c	1.49	0.65	0.00	3.00
T3_d	1.75	0.57	1.00	3.00
TOTAL	4.93	1.24	2.76	7.16

N = 34, Note: Test 1 (T1), Test 2 (T2), Test 3 (T3).

Question “a” has a weight of 1.0 and obtained means above 0.60 in all three tests, with the highest mean observed in T2 ($M = 0.89$, $SD = 0.15$). In contrast, question “b” holds a weight of 2.0, and the highest mean response also occurred in T2 ($M = 0.97$, $SD = 0.38$). However, all three means were much lower when compared to question “a”. Question “c” has a weight of 3.0 and attained its highest mean response at T1 ($M = 1.67$, $SD = 0.79$). Nevertheless, the means for question “c” remained lower compared to the other questions. Finally, question “d”, with a weight of 4.0, obtained its highest mean response in T1 ($M = 1.96$, $SD = 0.88$). Similar to questions “b” and “c,” the means for question “d” were notably low and consistent across all three tests.

Table 2 also shows the total score for each test, with a maximum possible score of 10.0. The mean scores across all three tests were relatively low, with the highest mean score observed in T1 ($M = 5.20$, $SD = 1.44$). The scores are also quite similar, which could be attributed to the similarity in case content and the questions addressed across all tests. In **Table 3**, the Wilcoxon test highlights the statistical difference between the medians of the total scores.

Table 3. Comparison between total scores of the SJT.

Comparison	Total Score	M	Mdn	SD	<i>z</i>	<i>p</i>
T1-T2	Total Score T1	5.20	4.98	1.44	432	0.122
	Total Score T2	4.84	4.72	1.47		
T2-T3	Total Score T2	4.84	4.72	1.47	310	0.718
	Total Score T3	4.93	5.06	1.24		
T3-T1	Total Score T3	4.93	5.06	1.24	275	0.252
	Total Score T1	5.20	4.98	1.44		

$N = 34$. $p < 0.05$. Note. Mean (M), Median (Mdn), Standard Deviation (SD), Test 1 (T1), Test 2 (T2), Test 3 (T3).

Results displayed in **Table 3** indicate that there was no significant difference in total scores among the three SJTs ($p < 0.05$). However, when comparing the scores of identical questions in the different SJTs (“a”, “b”, “c” and “d”), the Wilcoxon test revealed significant differences in five out of twelve combinations: 1) T1_a (Mdn = 0.660) and T2_a (Mdn = 2.00), ($Z = 43.5$, $p \leq 0.001$); 2) T3_a (Mdn = 1.00) and T1_a (Mdn 0.660), ($Z = 442.0$; $p \leq 0.001$); 3) T2_b (Mdn = 1.06) and T3_b (Mdn = 1.06), ($Z = 172.0$, $p = 0.012$); 4) T1_c (Mdn = 2.00) and T2_c (Mdn = 1.00), ($Z = 265.5$, $p \leq 0.001$); and 5) T2_c (Mdn = 1.00) and T3_c (Mdn = 1.60), ($Z = 92.0$, $p = 0.034$). These findings, supported by median values and inferential tests, suggest that these five combinations of questions differ in terms of complexity levels.

5. Discussion

This study presented the process of construction and evaluation (semantic and content) of three SJTs focused on social skills within the collaborative work of

health care residents. The tests described different real-life situations commonly experienced by residents in Brazilian university hospitals, requiring participants to provide problem-solving responses by answering four open-ended questions. SJTs are recommended by health researchers for realistically simulating clinical practice dilemmas in multidisciplinary social interactions (Kiessling et al., 2016; Weng et al., 2018; Wolcott et al., 2020). Test content in this study was submitted to content analysis by professionals from different healthcare backgrounds to evaluate technical pertinence and applicability in professional practice.

The 15 interviews aimed at assessing the learning needs of medical and non-medical residents, as well as the content analyses conducted by 17 professionals from nine different healthcare academic backgrounds, played a key role in mitigating potential researcher bias and subjectivity when drafting the situational challenges featured in the tests. Additionally, this approach helped address the deficiency in educational interventions targeting social skills within healthcare teams (Anderson et al., 2017; Collins et al., 2021; Mata et al., 2021; Marlow et al., 2018; Patterson et al., 2016). The strategy of data collection and analysis through multiple sources also contributed to encompassing a broad spectrum of real-world elements across various healthcare settings (e.g. ICUs, outpatient clinics, inpatient units) and the crosscutting social skills essential for health professionals operating within a multi-professional context (Kiessling et al., 2016; Weng et al., 2018).

The professionals who participated in the content analysis stage exhibited a notable diversity in educational background, years of experience, geographical location (federative unit), and type of healthcare institution they were affiliated with (public or private). In instances where consensus was lacking on certain aspects of the content, adjustments were made, or the specific case was excluded, as illustrated by case 5. The diverse profiles of health care professionals significantly enrich the depth and comprehensiveness of construct analysis by incorporating a broad spectrum of perspectives and opinions. Similarly, the professionals involved in the pilot application demonstrated heterogeneity, including both experts and non-experts. This diversity allowed for a substantial range of both correct and incorrect responses (Cechella et al., 2021; Christian et al., 2010; Cualheta & Abbad, 2021).

This study aims to enhance the reliability and validity of learning assessment in the fields of education and healthcare using SJTs that are based on realistic scenarios and observable behaviours exhibited by professionals in clinical settings. One of the contributions of this study is the presentation of alternatives to the prevalent use of self-assessments and subjective assessments within the field of health education (Barron et al., 2022; Havyer et al., 2016). SJTs differ from objective tests that measure knowledge based on closed-ended responses, such as true-false, gap-filling, multiple choice, and similar formats. This distinction is particularly evident as SJTs require open-ended answers regarding behaviours, skills, and attitudes, all inspired by cases derived from practical scenarios.

The study addresses a recognized demand within the literature, which underscores the importance of integrating not only technical-scientific skills but also social skills within healthcare training curricula, particularly for residency teams (Abbad et al., 2016; Lamba et al., 2016). The scenarios described in the SJTs closely mirror real-world situations and may serve as educational resources within active teaching methodologies aimed at imparting practical skills to multi-professional teams (Cox et al., 2017).

The tests prepared in this study (Appendix A) show examples of work situations involving professionals of different formations that can be used intentionally as a didactic resource to improve the learning of medical and non-medical residents during multi-professional meetings. Tests can also be used to evaluate learning outcomes in practical training that simulate the debate between medical and non-medical residents and the joint resolution of technical and behavioral problems.

The STJs were built using critical incidents extracted from a learning needs assessment, serving as the foundation for both their development and the strategic planning of educational interventions targeting specific social skills relevant to the multi-professional context of health residents. The alignment between diagnosing learning demands, planning training sessions, and formulating assessment instruments is pivotal in fostering effective learning and the practical application of acquired social skills in the workplace (Aguinis & Kraiger, 2009; Bell et al., 2017).

Test items were presented to participants in an online written format. However, it should be noted that SJTs can take various formats, such as video presentations, simulations, or role-plays. The language used in these tests was compatible with the socio-professional backgrounds of the participants. (Christian et al., 2010; Reed et al., 2022; Patterson et al., 2016; Smith et al., 2020). Regardless of the chosen format, scenarios in tests should be collaboratively developed with experts on the subject to ensure technical fidelity to the reality under investigation. The current study adhered to this recommendation and followed essential steps outlined in the scientific literature for designing high-quality SJTs. These steps included the establishment of key responses and scoring methods, the conduct of pilot testing, and psychometric analyses (Cechella et al., 2021; Cua-lheta & Abbad, 2021; Patterson et al., 2016).

The formulation of the four open-ended questions was based on the instructional objectives established for the social skills training of multi-professional healthcare residents. These instructional objectives were also aligned with the social skills deficits identified in the interviews. The definition of accurate instructional objectives, articulated as observable performances and aligned with learning needs, contributes to the construction of learning items that exhibit both consistency and reliability in assessing their intended constructs (Bell et al., 2017; Christian et al., 2010).

The levels of complexity proposed in the four open-ended questions, accord-

ing to Bloom's taxonomy (1956), are consistent with the complexity inherent in a SJT, since this type of assessment typically involves complex situations that require multifaceted problem-solving approaches. Therefore, a pilot test was administered, and a scoring worksheet was prepared to investigate a broad spectrum of potential responses from different professionals (Cechella et al., 2021; Cualheta & Abbad, 2021).

5.1. Limitations

Despite the contributions of this study, it is essential to acknowledge the challenges associated with accessing health professionals, which stem from the inherent nature of the research field. The relatively small number of participants may be attributed to the long time required to complete the three SJTs. Each test encompassed a case description accompanied by four open-ended questions that required a range of cognitive abilities, spanning from comprehension (recollection of facts) to creative thinking (proposing solutions). A significant number of health professionals were unable to adhere to the research due to the limited timeframe available to answer the 12 open-ended questions. Kang et al. (Kang et al., 2022) also discussed the practicality and feasibility of employing these tests in the healthcare context. The authors pointed out that the heavy workload of these professionals, coupled with the scarcity of free time available to complete the survey, might negatively influence the quality of responses when applying such instruments in the healthcare setting.

Another limitation refers to the degree of complexity of some items, particularly those consisting of open-ended questions. While all tests addressed the three social skills indicated as learning gaps among residents, the situations described were different, originating from different healthcare contexts such as the ICU, outpatient clinics, and inpatient units. The inherent characteristics of each context could have contributed to the heterogeneity of responses, even when posed with identical questions. Additionally, the equivalence of content complexity was not consistently maintained across the 5 combinations of items in different versions of the three SJTs (T1_a and T2_a, T3_a and T1_a, T2_b and T3_b, T1_c and T2_c, T2_c and T3_c). Although the cases depicted the same social skills and maintained a similar character count, the test items/questions may have been sensitive to specific characteristics of contexts within the healthcare field. The work environment and culture within private and public hospitals often differ significantly, potentially impacting socio-affective interactions and collaborative dynamics among multiprofessional teams.

5.2. Future Research

Further research is suggested to assess the efficacy of SJT-based learning in acquiring social skills relevant to collaborative multiprofessional healthcare. These skills should be identified in systematic assessments of learning needs conducted with professionals operating within the specific investigated scenario. The aim is

to build SJTs that encompass content aligned with the authentic work context. The SJTs developed in this study primarily focused on selected sectors of a hospital, including the Intensive Care Unit (ICU), and outpatient and inpatient units. It is advisable to develop SJTs that are sensitive to other healthcare settings (e.g., primary care, emergency care, surgical centres, etc.) and other sectors of society, such as primary education, given that the acquisition of social skills can and should occur throughout one's lifespan (Del Prette & Del Prette, 2018). In addition, future studies could evaluate whether the written responses match the behaviours shown in simulations.

When built with methodological rigour and aligned with the specific educational needs of a target audience, SJTs may be incorporated into interprofessional education initiatives within the healthcare domain. This integration contributes to a learning process that is attuned to the realities faced by both students and professionals (Wolcott et al., 2020). SJTs can serve as instruments for integrating training in a formative manner throughout a capacity-building project, in the form of exercises that require problem-solving skills (Cox et al., 2017), or as a summative component after the capacity-building process. Further studies are then suggested to present the results of the application of these SJTs in different contexts and at different moments, both before and after training programs.

It is relevant to develop more than one version of the same SJT, as this allows for its application in studies with robust designs, including pre- and post-tests, which are recommended by the scientific literature. The complexity of the test content can also be explored in additional studies that employ experimental and quasi-experimental designs. This approach contributes to the assessment of learning effects within groups and between different groups. Another recommendation for future studies involves investigating the correlation between the acquisition of social skills from TD&E programs and team-based outcomes, including metrics such as customer satisfaction ratings and other pertinent healthcare indicators.

6. Final Considerations

The current study contributes to the body of knowledge regarding the development of SJTs specifically tailored for the evaluation of social skills acquisition among cohorts of both medical and non-medical residents. Investment in the development of training and assessment instruments focused on non-technical skills plays a key role in mitigating difficulties in the socio-affective interaction of health care professionals. Such shortcomings, if not addressed, have the potential to negatively influence the quality of patient care. Disabled communication between professionals in a multi-professional team can contribute to the increase in medication and prescription errors, care delays, forgetfulness, stress, and illness of professionals, among other factors harmful to patient safety and care. (Jin et al., 2022; Lamba et al., 2016; Riley et al., 2011, Shapiro et al., 2004).

SJTs emerge as a valuable and versatile tool, applicable not only within the formative phase of resident training but also within ongoing healthcare education initiatives, including undergraduate programs that may currently still offer limited opportunities for the acquisition of practical, situational, and interprofessional socio-affective skills. These SJTs are expected to be continuously applied for various purposes, such as the assessment of training needs, the formulation of instructional materials, preliminary skill appraisal prior to training, and subsequent evaluation post-training.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

- Abbad, G. da S., Parreira, C., Pinho, D., Queiroz, E., Torres, A., Furlanetto, D., Jorge, A., & Silva, N. (2016). *Formação e Processos Educativos em Saúde. Ensino na Saúde no Brasil* (27-48). Juruá.
- Aguinis, H., & Kraiger, K. (2009). Benefits of Training and Development for Individuals and Teams, Organizations, and Society. *Annual Review of Psychology*, *60*, 451-474. <https://doi.org/10.1146/annurev.psych.60.110707.163505>
- Anderson, R., Thier, M., & Pitts, C. (2017). Interpersonal and Intrapersonal Skill Assessment Alternatives: Self-Reports, Situational-Judgment Tests, and Discrete-Choice Experiments. *Learning and Individual Differences*, *53*, 47-60. <https://doi.org/10.1016/j.lindif.2016.10.017>
- Barron, L. G., Ogle, A. D., & Rowe, K. (2022). Improving the Effectiveness of Embedded Behavioral Health Personnel through Situational Judgment Training. *Military Psychology*, *34*, 377-387. <https://doi.org/10.1080/08995605.2021.1971938>
- Bell, B. S., Tannenbaum, S. I., Ford, J. K., Noe, R. A., & Kraiger, K. (2017). 100 Years of Training and Development Research: What We Know and Where We Should Go. *Journal of Applied Psychology*, *102*, 305-323. <https://doi.org/10.1037/apl0000142>
- Cechella, F., Abbad, G., & Wagner, R. (2021). Leveraging Learning with Gamification: An Experimental Case Study with Bank Managers. *Computers in Human Behavior Reports*, *3*, Article 100044. <https://doi.org/10.1016/j.chbr.2020.100044>
- Chang, Y. C., Chou, L. T., Lin, H. L., Huang, S. F., Shih, M. C., Wu, M. C., Wu, C. L., Chen, P. T., & Chaou, C. H. (2019). An Interprofessional Training Program for Intra-hospital Transport of Critically Ill Patients: Model Build-Up and Assessment. *Journal of Interprofessional Care*. <https://doi.org/10.1080/13561820.2018.1560247>
- Christian, M. S., Edwards, B. D., & Bradley, J. C. (2010). Situational Judgment Tests: Constructs Assessed and a Meta-Analysis of Their Criterion-Related Validities. *Personnel Psychology*, *63*, 83-117. <https://doi.org/10.1111/j.1744-6570.2009.01163.x>
- Collins, L., Sicks, S., Hass, R. W., Vause-Earland, T., Ward, J., Newsome, C., & Khan, M. (2021). Self-Efficacy and Empathy Development through Interprofessional Student Hotspotting. *Journal of Interprofessional Care*, *35*, 320-323. <https://doi.org/10.1080/13561820.2020.1712337>
- Cox, C. B., Barron, L. G., Davis, W., & de la Garza, B. (2017). Using Situational Judgment Tests (SJTs) in Training: Development and Evaluation of a Structured, Low-Fidelity Scenario-Based Training Method. *Personnel Review*, *46*, 36-45.

<https://doi.org/10.1108/PR-05-2015-0137>

- Creswell, J. W., & Creswell, J. D. (2021). *Projeto de pesquisa: Métodos qualitativo, quantitativo e misto*. Penso Editora.
- Cualheta, L. P., & Abbad, G. D. S. (2021). Assessing Entrepreneurship Education Outcomes in an Innovative Way: Situational Judgment Tests. *Entrepreneurship Education and Pedagogy*, 5, 89-112. <https://doi.org/10.1177/2515127420975176>
- Del Prette, A., & Del Prette, Z. A. P. (2018). *Competência Social e Habilidades Sociais. Manual Teórico-Prático*. Editora Vozes.
- Havyer, R. D., Nelson, D. R., Wingo, M. T., Comfere, N. I., Halvorsen, A. J., McDonald, F. S., & Reed, D. A. (2016). Addressing the Interprofessional Collaboration Competencies of the Association of American Medical Colleges: A Systematic Review of Assessment Instruments in Undergraduate Medical Education. *Academic Medicine*, 91, 865-888. <https://doi.org/10.1097/ACM.0000000000001053>
- Jin, J., Son, Y. J., Tate, J. A., & Choi, J. (2022). Challenges and Learning Needs of Nurse-Patients' Family Communication: Focus Group Interviews with Intensive Care Unit Nurses in South Korea. *Evaluation & the Health Professions*, 45, 411-419. <https://doi.org/10.1177/01632787221076911>
- Kang, H., Flores-Sandoval, C., Law, B., & Sibbald, S. (2022). Interdisciplinary Health Care Evaluation Instruments: A Review of Psychometric Evidence. *Evaluation & the Health Professions*, 45, 223-234. <https://doi.org/10.1177/01632787211040859>
- Kiessling, C., Bauer, J., Gartmeier, M., Iblher, P., Karsten, G., Kiesewetter, J., Moeller, G., Wiesbeck, A., Zupanic, M., & Fisher, M. (2016). Development and Validation of a Computer-Based Situational Judgement Test to Assess Medical Students' Communication Skills in the Field of Shared Decision Making. *Patient Education and Counseling*, 99, 1858-1864. <https://doi.org/10.1016/j.pec.2016.06.006>
- Lamba, S., Tyrie, L. S., Bryczkowski, S., & Nagurka, R. (2016). Teaching Surgery Residents the Skills to Communicate Difficult News to Patient and Family Members: A Literature Review. *Journal of Palliative Medicine*, 19, 101-107. <https://doi.org/10.1089/jpm.2015.0292>
- Levitt, H. M., Bamberg, M., Creswell, J. W., Frost, D. M., Josselson, R., & Suárez-Orozco, C. (2018). Journal Article Reporting Standards for Qualitative Primary, Qualitative Meta-Analytic, and Mixed Methods Research in Psychology: The APA Publications and Communications Board Task Force Report. *American Psychologist*, 73, 26-46. <https://doi.org/10.1037/amp0000151>
- Lievens, F., & Sackett, P. R. (2012). The Validity of Interpersonal Skills Assessment via Situational Judgment Tests for Predicting Academic Success and Job Performance. *Journal of Applied Psychology*, 97, 460-468. <https://doi.org/10.1037/a0025741>
- Lineberry, M., Bryan, E., Brush, T., Carolan, T. F., Holness, D., Salas, E., & King, H. (2013). Measurement and Training of TeamSTEPPS® Dimensions Using the Medical Team Performance Assessment Tool. *Joint Commission Journal on Quality and Patient Safety*, 39, AP1-AP3. [https://doi.org/10.1016/S1553-7250\(13\)39013-8](https://doi.org/10.1016/S1553-7250(13)39013-8)
- Mahler, C., Berger, S., & Reeves, S. (2015). The Readiness for Interprofessional Learning Scale (RIPLS): A Problematic Evaluative Scale for the Interprofessional Field. *Journal of Interprofessional Care*, 29, 289-291. <https://doi.org/10.3109/13561820.2015.1059652>
- Marlow, S., Bisbey, T., Lacerenza, C., & Salas, E. (2018). Performance Measures for Health Care Teams: A Review. *Small Group Research*, 49, 306-356. <https://doi.org/10.1177/1046496417748196>
- Mata, Á. N. D. S., de Azevedo, K. P. M., Braga, L. P., de Medeiros, G. C. B. S., de Oliveira Segundo, V. H., Bezerra, I. N. M., Pimenta, I. D. S. F., Nicolás, I. M., & Piuevzam, G.

- (2021). Training in Communication Skills for Self-Efficacy of Health Professionals: A Systematic Review. *Human Resources for Health*, 19, Article No. 30. <https://doi.org/10.1186/s12960-021-00574-3>
- Patterson, F., Zibarras, L., & Ashworth, V. (2016). Situational Judgement Tests in Medical Education and Training: Research, Theory and Practice: AMEE Guide No. 100. *Medical Teacher*, 38, 3-17. <https://doi.org/10.3109/0142159X.2015.1072619>
- Peduzzi, M., Norman, I., Coster, S., & Meireles, E. (2015). Adaptação transcultural e validação da Readiness for Interprofessional Learning Scale no Brasil. *Revista da Escola de Enfermagem da USP*, 49, 7-15. <https://doi.org/10.1590/S0080-623420150000800002>
- Reed, B. N., Smith, K. J., Robinson, J. D., Haines, S. T., & Farland, M. Z. (2022). Situational Judgment Tests: An Introduction for Clinician Educators. *Journal of the American College of Clinical Pharmacy*, 5, 67-74. <https://doi.org/10.1002/jac5.1571>
- Reeves, S., Perrier, L., Goldman, J., Freeth, D., & Zwarenstein, M. (2013). Interprofessional Education: Effects on Professional Practice and Healthcare Outcomes. *Cochrane Database of Systematic Reviews*, No. 3, CD002213. <https://doi.org/10.1002/14651858.CD002213.pub3>
- Riley, W., Davis, S., Miller, K., Hansen, H., Sainfort, F., & Sweet, R. (2011). Didactic and Simulation Nontechnical Skills Team Training to Improve Perinatal Patient Outcomes in a Community Hospital. *The Joint Commission Journal on Quality and Patient Safety*, 37, 357-364. [https://doi.org/10.1016/S1553-7250\(11\)37046-8](https://doi.org/10.1016/S1553-7250(11)37046-8)
- Shapiro, M. J., Morey, J. C., Small, S. D., Langford, V., Kaylor, C. J., Jagminas, L., Suner, S., Salisbury, M. L., Simon, R., & Jay, G. D. (2004). Simulation Based Teamwork Training for Emergency Department Staff: Does It Improve Clinical Team Performance When Added to an Existing Didactic Teamwork Curriculum? *BMJ Quality & Safety*, 13, 417-421. <https://doi.org/10.1136/qshc.2003.005447>
- Shrader, S., Farland, M. Z., Danielson, J., Sicat, B., & Umland, E. M. (2017). A Systematic Review of Assessment Tools Measuring Interprofessional Education Outcomes Relevant to Pharmacy Education. *American Journal of Pharmaceutical Education*, 81, 119. <https://doi.org/10.5688/ajpe816119>
- Smith, K. J., Flaxman, C., Farland, M. Z., Thomas, A., Buring, S. M., Whalen, K., & Patterson, F. (2020). Development and Validation of a Situational Judgement Test to Assess Professionalism. *American Journal of Pharmaceutical Education*, 84, AJPE7771. <https://doi.org/10.5688/ajpe7771>
- Tabachnick, B. G., & Fidell, L. S. (2013). *Using Multivariate Statistics*. Pearson.
- Thistlethwaite, J., Moran, M., & World Health Organization Study Group on Interprofessional Education and Collaborative Practice (2010). Learning Outcomes for Interprofessional Education (IPE): Literature Review and Synthesis. *Journal of Interprofessional Care*, 24, 503-513. <https://doi.org/10.3109/13561820.2010.483366>
- Weng, Q. D., Yang, H., Lievens, F., & McDaniel, M. A. (2018). Optimizing the Validity of Situational Judgment Tests: The Importance of Scoring Methods. *Journal of Vocational Behavior*, 104, 199-209. <https://doi.org/10.1016/j.jvb.2017.11.005>
- Wolcott, M. D., Lobczowski, N. G., Zeeman, J. M., & McLaughlin, J. E. (2020). Situational Judgment Test Validity: An Exploratory Model of the Participant Response Process Using Cognitive and Think-Aloud Interviews. *BMC Medical Education*, 20, Article No. 506. <https://doi.org/10.1186/s12909-020-02410-z>

Appendixes

Appendix A. Situational Judgement Tests

Situational Judgement Test 1 (T1)

Situation: Non-aspiration of a patient's airway

You are a member of a team at the University Hospital and are attending a multiprofessional meeting at the Intensive Care Unit (ICU). You are still getting to know this team as you have been working with them for less than a month. In the initial team meeting, several topics are discussed. However, one particular issue, concerning patient Roseana, who is under sedation and mechanical ventilation, creates tension between two members: Luciana (nurse) and Felipe (physiotherapist). In the ICU, specific moments are designated for airway aspiration based on each patient's clinical condition. However, uncertainties and conflicts arose regarding the professional responsible for performing this procedure. Nurse Luciana, on duty at the time, did not perform the aspiration procedure on the patient who had excessive secretion. Furthermore, she documented on the patient's medical records that the procedure was not performed due to the absence of a physiotherapist in the unit, asserting that it was the physiotherapist's responsibility. She also registered that she observed an increase in the patient Roseana's respiratory effort, stating that she spent many hours without being aspirated. During the subsequent team meeting, physiotherapist Felipe, who was mentioned in the medical record, asked nurse Luciana about the record:

- *You are already here on duty, taking care of the patient. Would it hurt to do it? I always perform this procedure, even when I am alone without any support from your side of the team. It shouldn't be that a nurse couldn't step in too.*

Luciana, visibly frustrated, raises her voice, stating:

- *I'm not doing something that is not my responsibility! This has always been a physiotherapy task.*

The discussion continues, with both professionals shifting the responsibility to the other. Felipe seizes the moment and reports another concern.

- *You nursing staff can never do anything. Just Yesterday, I noticed a lung ventilator missing in one of the beds. I asked the nursing technician, who was assisting patient Roseana, and he bluntly claimed that it's the physiotherapist's responsibility to bring the ventilator to the patient's bed.*

- *And that's true! We already do a lot,* emphasizes Luciana.

The rest of the team remains silent, avoiding taking a stance as the debate continues between the two colleagues. You recall other conflicts between team members last week, which led to customer complaints filed with the Hospital ombudsman. At a given moment, you, as a team member, consider intervening to somehow steer the conversation away. However, they delve into a technical dialogue using specific language about the boundaries of their professions, making it difficult to understand the content of the discussion.

You contemplate how to address the issue without causing further embarrassment in the team meeting. How would you proceed in this situation? Answer the questions below, providing details on how you would handle it.

Answer the following questions based on the case report above:

- 1) What are the professions of the team members who interacted with the patient Roseana?
- 2) Describe three behaviours manifested by the team professionals that led to the delay in aspirating patient Roseana's airway.
- 3) Elaborate on how the team members could have acted to prevent situations similar to that of patient Roseana.
- 4) In light of the proposed solutions, how would you communicate with both professionals to prevent situations similar to that of patient Roseana? Describe what you would say to the team and how the dialogue might unfold.

Situational Judgement Test 2 (T2)

Situation: Patient is discharged without instructions from Occupational Therapy

You are a member of the health team at a University Hospital's Inpatient Unit. Juliana, an Occupational Therapist, gets to the hospital for her usual workday. Her initial task is to join the changeover meeting. During the meeting, upon listening to the final reports from the professionals who worked in the previous shift, she notices that a patient undergoing stroke rehabilitation has been discharged without her knowledge. In response, she approaches Nurse Similarly, who is sitting at the computer in the team room, and asks:

- *Hey Similarly! Good morning! Patient Kelly Regina, bed 15, should have been instructed about some procedures necessary for her recovery at home. Why wasn't I consulted? It's the third time this month that it happens. It's hard to work like this, huh?*

Similarly lowers her head, expressing regret for what happened and answers:

- *Juliana, I guess Dr. Diogo was in a hurry and couldn't find you at the moment of the discharge. He also commented that the most important instructions had already been given.*

Physiotherapist Lisa, who was also in the ward, overhears the dialogue and offers her input without looking at her two colleagues:

- *I was there at the time of discharge with Dr. Diogo. We never manage to find you, Juliana. The patient was anxious to go home.*

We have discussed about this patient extensively and you never mentioned anything about the need for such instructions. We couldn't have guessed it, right?

- Lisa, I always express my opinions during meetings, but the team barely consults me by the time of discharge. As a physiotherapist, you should know that this OT instruction is fundamental and crucial for the patient's recovery. Why didn't you try to contact me?

- Juliana, I think you should solve this with Dr. Diogo, okay? I don't want to be part of this mess. Good morning!

Juliana asks the team about the reason for the decisions, but the members cannot identify what caused the lack of communication. Then, one of the patients at the outpatient clinic seeks assistance, leading to the team members dispersing for their respective duties.

As an attentive team member who witnessed the whole discussion, you send a message on WhatsApp proposing a team meeting to discuss and solve the case. Despite everyone viewing the message, there is no response from any team member.

Answer the questions based on the case report above:

- 1) What are the professions of the team members who engaged with the patient Kelly Regina?
- 2) Describe three behaviours exhibited by the team professionals that led to Kelly Regina leaving the Hospital without receiving instructions from the Occupational Therapist Juliana.
- 3) Elaborate on the proactive measures the team professionals could have taken to prevent situations similar to that of Kelly Regina.
- 4) Given the aforementioned solutions, how would you communicate with your colleagues to avoid the recurrence of situations similar to that of Kelly Regina's? Describe what you would say to the team and how the conversation would look like.

Situational Judgement Test 3 (T3)

Situation: Change of the patient's diet

You are a member of the health team in the Inpatient Unit of a University Hospital. Leticia, a nutritionist, arrives at work visibly upset and requests a discussion with her team members. You are in the room, along with Nurse Eliane and Psychologist Gabriela. Leticia initiates the conversation:

- Hi people, good morning! I have assessed patient Luana in bed 7 and proposed a diet with specific quantities. I left everything written down for Dr. George to prescribe, along with instructions for the nursing staff to administer other medications. However, hours later, I found out that the values of the prescribed diet were completely different from my records. Can anyone help me understand what happened?

Nurse Eliane immediately checks the electronic record and answers:

- Hmm... Patient Luana? It was Paola, the new nurse, who administered the infusion. You should ask her. What I know is that the patient had diarrhoea and the diet was stopped this morning by Dr. George.

- How so? Besides administering the wrong amount of the diet, it's stopped without even discussing it with me? - asks Nutritionist Leticia.

The psychologist Gabriella, who was in the room and also followed Luana's case, shares her opinion with some embarrassment:

- Leticia, this patient's sister is also a nutritionist. She has been resistant and, jointly with her sister, she disagreed with the procedures adopted, including the nutritional values that you have been proposing. She spoke to Dr. George. So, he probably changed the values and did not inform you.

Leticia is taken aback by the information from Gabriela, and turns to nurse Eliane, who shakes her head confirming what the psychologist just said, and then comments:

- So, the nursing part is explained too. Paola (new nurse) probably followed the amended medical prescription when administering the diet. But she should have communicated with you too, Leticia.

Dissatisfied, Leticia turns to her teammates in the Unit, questioning why she always seems to be the last one informed. The team members are left without an answer. Frustrated, the nutritionist storms out of the room, expressing her intent to report the team to the Hospital's board of directors due to lack of communication among team members, which could jeopardize the patient's safety and health.

You rise from your seat, encouraging Leticia to return to the room and talk to the team.

Answer the questions based on the case report above:

- 1) What are the professions of the team members who interacted with the patient Luana?
- 2) Describe three behaviours manifested by the team members that contributed to the change in Luana's diet.
- 3) Explain how the team members could have acted to prevent situations similar to that of patient Luana.
- 4) Given the proposed solutions, how would you communicate with your colleagues to prevent the recurrence of situations similar to that of Luana's? Describe what you would say to the team and how the conversation would look like.

Source: Prepared by the authors.

Appendix B. Situational Judgment Tests Designed According to Learning Goals and Response Expectations (Cases 2, 4 and 6 - T1, T2 and T3)

Code	Item	Instructional Objective	Level of complexity	Expectation of response
C02A	Based on the case description, what are the professions of the team members who interacted with patient Roseana and failed to perform bronchoaspiration?	- To identify the socio-affective interactions generated between members of the multiprofessional team in everyday work.	Knowledge	To list specifically the 3 professions mentioned in the case: Physiotherapist, Nurse and Nursing Technician.
C02B	Describe three behaviours exhibited by the team professionals that may have contributed to the delay in aspirating the airways of patient Roseana.	- To distinguish tasks, responsibilities, and roles of each member of the multiprofessional healthcare team. - To support collaborative work in a multiprofessional team as a resource for improving patient care and safety.	Comprehension Acknowledgement	1) Misunderstanding of the roles and responsibilities of physiotherapists, nurses and nursing technicians. 2) Lack of communication among team members. 3) Individualized decision-making in patient care.
C02C	Explain how the team professionals could have acted to avoid situations like that of patient Roseana.	- To discuss failures in socio-affective interactions at work and their impacts on patient care.	Organization and Synthesis	The nurse should have aspirated the patient even in the absence of the physiotherapist. After aspirating the patient, the nurse should then have discussed with the team to reinforce agreements and/or reach a multiprofessional agreement. - Establish an appropriate time and location to meet with the entire team to address the issue.
C02D	Considering the proposed solutions, how would you communicate with the team members to prevent new cases similar to that of patient Roseana from happening?	- Propose solutions based on discussions with members of the multi-professional team, according to technical recommendations. - Decide on the care provided to the patient, in collaboration with the members of the multi-professional health team.	Organization and Assessment	- Emphasize the prioritization of patient care and avoid pointing fingers at specific individuals. The patient is the responsibility of the entire team! - Emphasize the need for collaborative work by the multi-professional team. - Collaborate with the team to establish new agreements and propose improvements for patient care.
Note. C02A: Question A, Case 2 (T1), C02B: Question B, Case 2 (T1), C02C: Question C, Case 2 (T1).				
C04A	According to the case description, what are the professions of the team members who interacted with patient Kelly Regina?	- Identify the socio-affective interactions generated among the members of the multiprofessional team in the daily work routine.	Knowledge	Listing precisely the 4 professions mentioned in the case: Occupational Therapy, Nursing, Physiotherapy, and Medicine.
C04B	Describe three behaviours exhibited by the team members that may have led to Kelly Regina being discharged without the guidance of Occupational Therapy.	- Distinguish tasks and responsibilities and roles of each member of the multi-professional team. - Support collaborative multi-professional teamwork as a resource for improving patient care and safety.	Understanding and Appreciation	1) Lack of understanding of the role of the occupational therapist in the discharge process 2) Lack of debate between team members. 3) Individualized decision-making in patient care.

Continued

C04C	Explain how the professionals on the team could have acted to avoid situations similar to that of Kelly Regina.	<ul style="list-style-type: none"> - Discussing failures in socio-affective interactions at work and their impacts on patient care. 	Organization and Synthesis	<p>Holding frequent meetings between team members to share information about the care provided to the patient. Informing the team when participation is impossible and possible referrals in these cases.</p> <ul style="list-style-type: none"> - Define an appropriate time and place to meet with the entire team to address the issue.
C04D	In view of the solutions presented, how would you talk to the professionals to prevent new cases like that of patient Kelly Regina from occurring?	<ul style="list-style-type: none"> - Propose solutions based on discussions with members of the multi-professional team, in accordance with technical recommendations. - Decide on the care provided to the patient, in collaboration with the members of the multi-professional health team. 	Organization and Evaluation	<ul style="list-style-type: none"> - Emphasize the prioritization of patient care and avoid pointing fingers at specific individuals. - Highlight the necessity of collaborative work within the multiprofessional team. - Collaborate with the team to establish new agreements and suggest improvements for patient care.

Note. C04B: Question B, Case 4 (T1), C04C: Question C, Case 4 (T2), C04D: Question D, Case 4 (T2).

C06A	According to the case description, what are the professions of the team members who interacted with patient Luana?	<ul style="list-style-type: none"> - Identify the socio-affective interactions generated among the members of the multiprofessional team in the daily work routine. 	Knowledge	To list exactly the 4 professions mentioned in the case: Nutrition, Psychology, Nursing and Medicine.
C06B	Describe three behaviours exhibited by the team members that may have led to the change in patient Luana's diet.	<ul style="list-style-type: none"> - Distinguish tasks and responsibilities and roles of each member of the multi-professional team. - Support collaborative multi-professional teamwork as a resource for improving patient care and safety. 	Understanding and Appreciation	<ol style="list-style-type: none"> 1) Misunderstanding of the role of the nutritionist. The new nurse also seemed not to know the team members and their potential internal work agreements. 2) Lack of debate between team members. 3) Individualized decision-making in patient care.
C06C	Explain how the professionals on the team could have acted to avoid situations like that of patient Luana.	<ul style="list-style-type: none"> - Discussing failures in socio-affective interactions at work and their impacts on patient care. 	Organization and Synthesis	<p>Internally aligning multiprofessional communication from the patient's sister's report about the nutritional values of the diet before the decision to change the diet. Socializing the new nurse with the adopted procedures and internal agreements of the multiprofessional team.</p> <ul style="list-style-type: none"> - Define the appropriate time and place to meet with the entire team to address the issue.
C06D	Considering the solutions presented, how would you communicate with the professionals to prevent new cases like that of patient Luana from happening?	<ul style="list-style-type: none"> - To propose solutions through discussions with members of the multiprofessional team, following technical recommendations. - To decide on the patient care collaboratively with members of the multiprofessional health team. 	Organization and Assessment	<ul style="list-style-type: none"> - Emphasize prioritizing patient care and avoid accusing specific individuals. The patient belongs to the whole team! - Emphasize the need for collaborative teamwork. - Make new agreements and suggestions for improving care together with the team.

Note. C06A: Question A from Case 6 (T3), C06B: Question C from Case 6 (T3), C06C: Question C from Case 6 (T3), C06D: Question D from Case 6 (T3).

Appendix C. Correction Spreadsheet

Item “a” - According to the case description what are the professions of the team members who interacted with the patient Roseana? VALUE: 1.0

Possible Responses	Note
Mentioning only physiotherapist, nurse and nursing technician.	1.0
Listing more than one health profession, including <u>the three</u> professions that interacted with the patient: physiotherapist, nurse, and nursing technician.	0.9
Listing more than one health profession, including <u>two</u> of the three professions that interacted with the patient.	0.7
Listing more than one health profession, including <u>one</u> of the professions that interacted with the patient.	0.5
Listing physiotherapist only	0.33
Mentioning nurse only	0.33
Mentioning nursing technician only	0.33
Mentioning another health professional only	0.2
Listing all those with technical training in health.	0
All are health professionals working in different areas.	0
I don't know	0
Blank	0

Source: Prepared by the authors.