

Interpersonal Skills Among Fourth- and Fifth-Year Medical Students

Validation of an Assessment Tool

Alexandre Morel, PhD Student;

Jean-Benoit Hardouin, PhD;

Pierre Pottier, PhD, MD

Objective: The aim of this study was to validate a French version of the standardized patient satisfaction questionnaire (SPSQ), an interpersonal skills assessment form, originally published in English, destined to be completed by SPs and used as a validated tool for the evaluation of communication skills of French medical students, during ambulatory care consultations.

Methods: During the academic year, fourth- and fifth-year medical students participated in mock ambulatory care consultations. The French SPSQ was completed by 16 SPs who performed scenarios written by professors of medicine. Validity evidence for test content, response process, internal structure, and relation to other variables was determined by calculating the Cronbach α coefficient, the Loewinger H index (response coherence), intraclass coefficients, the differences between fourth- and fifth-year students and by using item correlations as well as the correlation with an independent validated assessment form. G theory was used to identify the sources of variations in the mean of SPSQ.

Results: Four hundred thirty-three students were evaluated for a period of 2 months and participated in a total of 1703 mock consultations. The Cronbach α coefficient (0.94) was superior to that of the English SPSQ. The Loewinger H index was good ranging from 0.65 to 0.70. The reproducibility among external observers (intraclass correlation coefficient = 0.67) and between external observers and the SPs (intraclass correlation coefficient = 0.79 and 0.78, respectively) was satisfactory. Results of the French SPSQ were positively and significantly correlated to the scores obtained via an independent, validated, questionnaire ($R = 0.72$) and showed a significant difference between fourth- and fifth-year students' scores ($P < 0.0001$). G coefficients were estimated at 0.64 and 0.52 for year 4 and year 5, respectively.

Conclusions: Validity evidence supports the use of French-version SPSQ scores to screen for students experiencing difficulty with communication.

(*Sim Healthcare* 13:88–95, 2018)

Key Words: Simulation, standardized patient, medical students.

Since the reform of 2013, French medical studies have been focused on the acquisition of skills. Medical students are frequently placed in a situation of clinical practice in the hospital environment. They are, however, rarely in a position to propose their own solutions to a patient's problems, and they work primarily on the collection of medical data. On a smaller scale, learning experiences outside of the hospital environment do exist, in particular internships in general medicine, which have become compulsory during the second phase of French medical studies, although here again students play essentially an observational role.

Simulations offer the possibility of giving a student full responsibility for a patient by modeling a consultation in an ambulatory care environment. It allows for the development and an initial evaluation of certain skills specific to this environment. An actor is trained to portray in an iterative and replicable

manner, and a patient according to a predefined scenario is designated by the term "standardized patient" (SP).

The following three major skills can be observed during an ambulatory care medical consultation: the aptitude to analyze a medical problem presented by a patient, the proposal of a solution (provide a diagnosis and/or a treatment), and the establishment and maintenance throughout the duration of the consultation of a quality doctor-patient relationship.

It is only since the beginning of 2000 that "standardized patient" programs have been developed in France, as a method of teaching and of evaluating clinical reasoning and doctor-patient relationships.

Our organization introduced this method in 2009 as a tool for teaching clinical reasoning and doctor-patient relationship skills and not as an *Objective Structure Clinical Examination* type of evaluation method. The SPs program was designed in keeping with current recommendations.¹ This includes curriculum integration, repetitive practice back and forth between real and simulated patients, sufficient time devoted to systematic feedback and debriefing, as well as that to SPs training.

Although difficult for an SP to empirically judge the accuracy of the medical reasoning behind a clinical intervention, he

From the UMR 1246 INSERM "methodS in Patient-centered outcomes & HEalth ResEarch" (A.M., J.B.H., P.P.), Université de Nantes; and Service de Médecine Interne (P.P.), CHU de Nantes, Nantes, France.

Reprints: Alexandre Morel, PhD Student, Université de Nantes Nantes, Pays de la Loire France (e-mail: alexandre.morel@univ-brest.fr).

The authors declare no conflict of interest.

Copyright © 2018 Society for Simulation in Healthcare

DOI: 10.1097/SIH.0000000000000308

can, nonetheless, form a pertinent opinion with respect to the quality of the doctor-patient relationship.

In our experience, SPs have been able, on a regular basis, to identify students experiencing difficulty with communication. Students themselves were not always aware of their own difficulties.

In light of the number of students experiencing difficulties, we considered the implementation of a remedial program similar to those in place in other faculties.

Life coaching² and life skills training,^{3,4} teaching strategies for coping with stress,⁵ and courses on relaxation techniques⁶ may improve the well-being and the performance of students, who face difficulties with both their attitudes and their communication with patients. However, these programs have financial implications, which limit their use to a select number of students.

It has therefore become necessary to rationalize the evaluation process of interpersonal skills done by SPs, to focus remedial intervention on the students most in need.

Unfortunately, most of the tools for the evaluation of interpersonal skills among medical students in an ambulatory care environment have only been developed in English except for the French version of the Maastricht questionnaire,⁷ which was not chosen as a reference tool because it seemed to comprise too many questions and to explore dimensions of the consultation, which were above the standard of our undergraduate students (such as ability to synthesize data and to explain diagnostic and therapeutic options).

These questionnaires were essentially designed to be used by teachers^{8–14} by authentic patients^{14–19} or by SPs.^{16–19} Some are addressed to practicing doctors, whereas others are aimed at first-phase medical students.

The physical presence of an observer during mock medical consultations can disturb students and alter the evaluation. It is also an expensive method that is time-consuming for teachers. Indirect evaluations using audiovisual recordings have been validated for both simulations¹⁸ and real-life situations.¹⁹ The recordings can be evaluated by patients themselves or by teachers.

We therefore aim to develop an assessment tool with validity evidence to detect students experiencing difficulty with doctor-patient communication.

According to Downing²⁰ and the Standards for Educational and Psychological Testing,²¹ there are five elements that support the validity of an evaluation tool as follows:

1. evidence based on test content, which explores the relationship between the content of the test and the construct that is intended to measure,
2. evidence regarding the response process, which refers to the integrity of data and how all sources of error associated with the test are controlled or eliminated,
3. evidence regarding internal structure, which relates to the psychometric characteristics of the questions in terms of reproducibility and generalizability,
4. the relation to other variables, exploring both the convergent and divergent correlations, and
5. the consequences addressing the impact on the examinees from assessment scores and the impact of assessments on teaching and learning.

The aims of this study were to implement a tool for detecting students experiencing difficulty with doctor-patient

communication and to establish its validity with the first four elements of evidence: test content, response process, internal structure, and relation to other variables.

A later study will aim to show the validity with respect to the fifth element of evidence, by assessing the impact that the remedial program has had on the communication skills of the students, recognized by the SP program, experiencing difficulty.

METHOD

This study was submitted to an independent local ethics committee and received approval in March 2015 under the number GNEDS-03.03.2015.

Educational Context of the Validation of the Questionnaire

This study was performed during the biannual SP sessions included in the clinical rotations. The entire classes of fourth- and fifth-year medical students (during the 2014–2015 academic year), except repeat students, were required to participate in this study after having been duly informed.

At the Nantes Medical Faculty, during an academic year, each fourth- and fifth-year student participates in two series of distinct mock ambulatory care consultations during the first and the second semester. Each series consists of four mock consultations of a duration of 15 minutes each. Students from a same class were assessed using the same set of scenarios. Students have their own medical office with an examination table and the necessary tools for a clinical examination. Each student receives mock patients from a designated waiting room. He or she then proceeds to do a complete clinical investigation and physical examination in accordance with the main medical complaint of the patient. Students do not have access to any additional information about their patient's pathology over and above what they collect during the medical interview and during the physical examination. At the end of the consultation, students must give a diagnosis and/or a therapeutic plan while having developed a quality doctor-patient relationship in a limited timeframe. At the close of every consultation, the SP filled out an online questionnaire of evaluation, which was followed by an individual debriefing done by the SP who had just played the role of mock patient. The encounters were not observed by faculty. Finally, at the conclusion of a series of four mock consultations, a debriefing in groups of 16 students was performed in the presence of one of the SPs, a university hospital doctor and an experienced host facilitator.

This study was implemented for a period of 6 weeks from February to March 2015 at the University of Nantes (France) corresponding to the second series of the fourth and fifth year. During this period, 1770 mock consultations were scheduled.

The evaluation of students' difficulties during these consultations was systematic. Although the results of the evaluations were considered as confidential, they were available to students who wished to know their results.

Evidence of Validity

Test Content

Specifications. In light of the objective of the questionnaire, we began by establishing the specifications of the ideal tool. The latter should be filled out by the SPs at the evaluation phase of an ambulatory consultation.

The items on the questionnaire, which evaluate interpersonal skills, must be compatible with the academic curriculum. This is in keeping with the first theme of the second cycle of French medical studies, which comprises a total of 387 themes, entitled “Doctor-Patient Relationships.” Furthermore, in our faculty, we defined the specific learning outcomes of this theme as inspired by the Calgary-Cambridge Guide.¹⁰

Moreover, the items of this tool require a high degree of thoroughness in relation to our learning guidelines, which includes 56 themes of which 44 are considered as undergraduate level and 12 as registrar level.

Literature Review. We subsequently undertook a systematic literature review, in search of tools compatible with our specifications. Our literature review was done through PubMed using the following key words: “students,” “assessment,” “interpersonal,” “communication,” and “skills.” To select our tool, we also relied on an existing article reporting a literature review.²²

Translation of the English Version of the Tool. The translation of the integral English version of the tool was done according to a sequential method. Firstly, an initial translation from English to French was undertaken. A translation of this French version back into English was then completed by a highly proficient native English speaker residing in the United Kingdom, ensuring the strict fidelity to the original English version in terms of concepts, items, and dimensions.

A third translation of the second English version was then completed by a highly proficient native English speaker, teaching English at the University of Nantes, but who had not participated in the first two phases.

The final version of the questionnaire was obtained after adjustment of the second French translation and by comparison of both English versions. In keeping with the objective, this process insured that the translation was not simply literal but rather true to the meaning of the original tool.

Scenario Content. We were assured that the themes and the educational objectives of the scenarios were perfectly compatible with the curriculum. We also insured that relative academic modules were completed before the mock consultations. The themes of the scenarios were type 1 diabetes, VIH infection, frontal sinusitis, and leg ulcer for the fourth-year students and chronic general aching, dementia, asthma, and depressive episode for the fifth-year students.

The scenarios were written by the professors of medicine responsible for the teaching of the theme of a given scenario. Each scenario was reviewed by an internist, expert in the field of teaching clinical reasoning.

In Support of the “Response Process”

All the SPs taking part in the study had experience as SPs and were experienced with filling out evaluation questionnaires at the end of each consultation. Completion of the evaluation questionnaires was done online, immediately after the consultation, which minimized the risk of missing data or a recall bias. The Standardized Patient Satisfaction Questionnaire (SPSQ) was filled out by 16 SPs who portrayed patients. All SPs had received previous training in the use of the questionnaire by an internist in charge of interpersonal skills

learning (1 hour of training per scenario). To optimize the standardization of the assessments, SPs training was inspired from the recommendations done by Cleland et al. in the AMEE Guide No. 42.²³ During this training, the purpose and the content of all items of the questionnaire were carefully explained and all questions of the SPs were answered, making sure that each term of the items was understood in the same way by all the SPs.

The concordance of the SPs scoring was verified by the correlations between the scores given by the SPs and the scores given by physicians from corresponding video recordings.

Thirty-one audiovisual recordings were randomly selected to guarantee a representative with respect to the different scenarios, SPs, and classes. Questionnaires were completed after visualization of the recordings by two observers, composed of a professor in medicine and the principal investigator of this study (who had scientific and not medical qualifications). The intraclass correlation coefficients (ICCs) were calculated between the observers to estimate the interevaluator reproducibility as well as between the observers and the SPs to estimate the reproducibility of the scoring done by the SPs. The reproducibility was considered as excellent when the ICC was greater than 0.8, good for coefficients between 0.6 and 0.8, and mediocre for coefficients between 0.4 and 0.6.²⁴

Internal Structure

To show the validity of the questionnaire, several statistical analyses were conducted.

The Internal Consistency of the Questionnaire. The internal consistency was estimated by calculating the standardized Cronbach α coefficient. This coefficient measures the adjusted (by the number of items) degree of interrelation between items on a questionnaire. A coefficient greater than 0.7 is usually considered as satisfactory, whereas a coefficient greater than 0.9 is considered to be excellent.

The Structural Validity. The tool was expected to have a unidimensional structure as it aims at measuring only a single concept (and obtaining a single score). This postulated unidimensionality can be verified by the degree of variance of the items defined by the first main component of an analysis of the main components. If the eigenvalue of the first principal component was the only one eigenvalue greater than 1, the unidimensionality of the questionnaire was considered to be satisfactory.²⁵ Furthermore, this unidimensionality was confirmed by the correlation coefficients between each item and the total score. The validity of a structure was considered satisfactory for coefficients greater than 0.4 and good for coefficients greater than 0.7.

The Response Coherence. It was measured via the calculation of Loevingers H coefficient. A response is defined as inconsistent when a failure is recorded after having succeeded when faced with a more difficult element. This incoherence is called Guttman error. The H index can be defined as a function of Guttman error. In practice, responses to items are generally considered as coherent if H is between 0.3 and 0.5 and highly coherent if H is in excess of 0.5. A high level of coherence of the responses argues for the unidimensionality of the questionnaire.

Generalizability Study (G study). A random-effect variance analysis has been performed and followed by a G study²⁶ using the sum of the squares to estimate the relative contribution of student's responses, cases, and SPs to the variance of the SPSQ mean, taking into account interactions between SPs and students and between SPs and cases. In this G analysis, students were considered as object of measurement (or differentiation facet), whereas cases and SPs were considered as facets of measurement (or instrumentation facets).

Our study was not fully cross-designed because one student could see only one time one SP and one case. Nevertheless, one SP could play several scenarios for different students but never met the same student. Consequently, all the interactions including the students' effect (students \times cases, students \times SPs, and students \times SPs \times cases) cannot be completely estimated because they required the estimations of a more important number of parameters (and so of degrees of freedom) than the number of available data. Therefore, for the application of the G theory, only single effects (students, cases, and SPs) and the interactions cases \times SPs can be totally estimated, and the rest of information cannot be distinguished between interactions including students.

G coefficient ranges from 0 to 1. As mentioned in article by Crossley et al,²⁶ "It provides a measure of how confident one can be that any differences detected between assesseees are real differences. This coefficient will always be lower than a classical reliability coefficient (interobserver, intraobserver, or test-retest reliability) because it takes into account of all possible sources of error at the same time. [...] $G = 0.8$ is the generally accepted threshold of reliability for high-stakes judgements."

Relationship With Other Variables

The Convergent Validity. The positive convergent validity was estimated via the correlation coefficient (Pearson correlation coefficient) between the score of the French version of the tool and another validated evaluation questionnaire of interpersonal skills: the Maastricht questionnaire.⁷

The Maastricht questionnaire allows for the evaluation of the advised medical care (4 questions), of communication skills at each phase of the consultation (6 questions), and of general interpersonal skills (6 questions). We expected a good correlation (superior to 0.7) with respect to the general communication score and a moderate correlation (superior to 0.4) with respect to the communication score per phase because SPSQ has not been designed to assess communication according to different phases of the consultation. It has been completed online by SPs simultaneously with SPSQ at the end of each simulated consultation.

Known Groups Validity. It was estimated via the study of the evolution of interpersonal skills between the fourth and fifth year. We expected a significant increase in the SPSQ scores between the fourth- and fifth-year students. This assessment was done using a Student *t* test.

Software

The statistical analyses were accomplished using the software Stata 13 (Stata Corp; 2013). The generalizability study has been performed using the software Edug 1.6 (Educant Inc).

RESULTS

Number of Participating Students and Number of Evaluated Consultations

Four hundred thirty-three students of the 443 initially planned participated in the mock consultation sessions. Ten students did not attend any of the mock consultations, equal to a total of 40 mock consultations that were not done. Twenty-five mock consultations could not be done because of a students' late arrival. Four assessments were deleted because of problems with student identification. Thus, among the 1770 mock consultations initially scheduled, a total of 1703 were evaluated for a period of 2 months (February–March 2015).

The mean numbers of encounters by SPs and by cases were 106.4 (38.8) and 212.9 (2.9), respectively. Fourth- and fifth-year students attended to 855 and 848 encounters, respectively.

Standardized Patient Satisfaction Questionnaire varied from 16.0 to 26.0 according to SPs rating and from 18.7 to 22.3 according to cases.

Evidence for Validity Test Content

Selection of a Tool Compatible With the Specifications. After an analysis of literature, the SPSQ^{22,27} seemed to be a relevant tool for the screening of interpersonal skills difficulties in medical students. Indeed, it has been used, in a similar educational context, in diagnostic medicine in the United States, during simulated sessions called *Objective Structured Clinical Examination*, completed by SPs at the end of sessions to evaluate interpersonal skills on a Likert scale of 10 items ranging from 1 to 5. The Cronbach α coefficient was 0.90. The SPSQ score was strongly correlated ($R = 0.75$) to the score of the evaluation of interpersonal skills given by an external observer during sessions of objective structured clinical examinations.

The items evaluated on this scale seem similar to those upon which we wished to evaluate our students.

Indeed, the comparison between the 10 items on the SPSQ and the curriculum of the first theme entitled "doctor-patient relationships" shows a 100% compatibility (none of the SPSQ items are off the topic) and a 68% coverage rate (30 criteria, of the 44 criteria which correspond to the undergraduate level on the Calgary-Cambridge guide, are covered by the SPSQ items).

French Version of the Questionnaire. The SPSQ was translated according to the procedure described in the method section (Appendix 1).

The French version of the SPSQ (Appendix 2) includes 10 questions, presented on a single page, and offers the possibility of producing a score on a five-dimension Likert scale between "mediocre" (poor) to "excellent" (excellent). The time required to complete the questionnaire was approximately 4 to 5 minutes. A "mediocre" evaluation was attributed a score of 0 whereas an "excellent" evaluation was given a score of 4; the quantitative equivalences of evaluations were not known by the SPs.

Response Process

Missing Data. No missing data were found.

Interevaluator Reproducibility. The ICCs were respectively 0.67 between the two observers, 0.79 between observer 1 and the SPs, and 0.78 between observer 2 and the SPs. The global ICC was estimated at 0.82.

Internal Structure

Internal Consistency. The Cronbach α coefficient was computed at 0.94, which is excellent.

To ensure that multiple assessments per student did not have a significant influence on the Cronbach α coefficient, we recalculated it after randomly selecting only one assessment per student ($n = 432$). No distinction was made between academic year, scenarios, and SPs. The recalculated Cronbach α coefficient was 0.95, and this shows that this test was not influenced by the repetition of assessments.

Structural Validity. The analysis of main components allows for the accentuation of a single dimension measured by all 10 items of the SPSQ. The eigenvalue of the first main component was markedly superior to 1 (6.65), whereas the eigenvalue of the second component was, indeed, inferior to 1 (0.91). The coefficients of correlation between each item and the total score of the SPSQ range from 0.76 (item 3) to 0.86 (item 10). These results validate the hypothesis that the SPSQ measures only a single construct.

Response Coherence. Loevinger H index was estimated at 0.70, with figures per item ranging from 0.65 (item 3) to 0.73 (item 10). The coherence between the items of the SPSQ was thus verified.

Generalizability Study. The variance due to the interaction between SPs and students represented the main part of the variance of the SPSQ mean, ie, 80% in the fourth-year students and 84% the fifth-year students. The contribution of cases to the variance was negligible. The G coefficients were 0.64 and 0.52 in the fourth year students and in the fifth-year students, respectively (Table 1).

Relationship With Other Variables

Convergent Validity. The correlation coefficients between the SPSQ score and the communication skills score of the Maastricht questionnaire were respectively 0.72 with the total score of communication and 0.67 with the communication score per phase. These values confirm the convergent validity between these two questionnaires.

Known Groups Validity. The score obtained by the fifth-year students was 21.1 (7.6) versus 19.7 (6.3) for the fourth-year students ($P < 0.0001$, Student t test).

DISCUSSION

The primary objective of this study was to present arguments for the validity of a French version of an interpersonal skills evaluation questionnaire aimed at screening students experiencing difficulties. The English version of the questionnaire was selected according to precise specifications.

TABLE 1. Variance Analysis and G Study

	Year 4		
	Sum of Squares	Degree of Freedom	% of Total Variance
Students	14,842	218	13.3
SPs	3506	10	4.5
cases	89	3	0.0
Interaction SPs/cases	285	10	1.7
All other interactions*	14,848	613	80.5
G coefficient		0.64	
	Year 5		
	Sum of Squares	Degree of Freedom	% of Total Variance
students	16,018	214	7.7
SPs	7680	11	6.9
cases	113	3	0
interaction SPs/cases	348	12	1.4
All other interactions*	21,634	607	84
G coefficient		0.52	

*Nondistinguishable in this model (insufficient degree of freedom).

Several results, obtained from a large number of mock consultations (1703), support the validity of the questionnaire:

- the internal consistency and the response coherence, measures of the questionnaires ability to produce a precise score to assess a single concept (in this case the doctor-patient relationship), were in our study highly accurate (0.94 Cronbach α coefficient, superior to that of the English version, and a Loevinger H index of 0.70),
- the ICC between external observers was good (0.67), demonstrating a satisfying reproducibility of the questionnaire when used to assess communication skills by faculty from videos recordings,
- the reproducibility between each observer and the SP was good because observer 1 and 2 had respectively an ICC of 0.79 and 0.78, respectively. These data demonstrate the good reliability of the questionnaire when used by SPs.

In terms of structure, this questionnaire aimed at studying a single dimension variable, in this case the quality of the doctor-patient relationship. This unidimensionality was verified within our cohort. The scores obtained with the SPSQ were positively and significantly correlated to the scores obtained with a validated questionnaire (the Maastricht questionnaire), used to estimate the clinical skills of doctors established in an ambulatory care environment by means of data collection, of clinical reasoning, of the quality of suggested medical care, and of the quality of communication. This good correlation further strengthens the validity of our French version of the SPSQ. The SPSQ is, however, different from the Maastricht questionnaire because it is easier to complete (10 items instead of 16) while being targeted toward interpersonal skills.

The questionnaire demonstrates a good discriminant validity because it is capable of detecting a significant difference between the scores obtained by fourth- and fifth-year students ($P < 0.0001$) although it still remains to be demonstrated whether or not this difference is educationally relevant (improvement of 1.6 points out of 40).

A relatively low global average score for the two years combined of 20/40 can be noted. This result can be explained by the fact that the questionnaire was initially designed to evaluate experienced doctors.

The high proportion of variance explained by the interaction between SPs and students supports the fact that even if the SPs training aimed at a high degree of standardization, each encounter between an SP and a student was unique leading to specific situations and variable results in SPSQ scores.

Study Limitations

Firstly, G coefficients (0.64 in fourth-year students and 0.52 in fifth-year students) showed that a nonnegligible part of the true variance of SPSQ was not explained by our conditions of test (4 different SPs and 4 different cases). Higher G coefficients might be obtained increasing the number of cases, which was relatively small as a screening test for student interpersonal difficulties.

Due to the fact that the validation of the questionnaire has been done in a simulated environment comprising SPs within a single center, the applicability of our results to other centers and to other uses, especially to an authentic context (with real patients), is still to be shown. Indeed, G study showed that applying the test to different classes of students, which implied different cases and SPs, modified the reliability of the test (G coefficients decreased from 0.64 to 0.52 between the fourth and the fifth year).

The SPs training for SPSQ rating was probably not long enough to ensure an optimal standardization. Indeed, Wallace et al²⁸ suggested that additional training should be required for assessment purposes. To ensure a better accuracy of SPs screening for students with interpersonal difficulties, it might be useful to extend our period of training from 1 to 3 hours and to add some practice exercises from video examples as recommended.

Our study provides little information on the cohort of participating students. However, as the entire year participated, the cohort was surely a true reflexion of the population within our institution. In France, fourth-year medical students are on average 22 years of age, which is younger than in some countries (notably North America). A lack of maturity could therefore explain lower scores of interpersonal skills.

The depth of the SPSQ with respect to certain aspects of the “doctor-patient relationships” theme (our blueprint) of the curriculum could be increased via the exploration of other important areas of communication such as the understanding of information provided, encouragement of the patient to express their feelings, asking for patients authorization before, and the explication of the clinical examination, regular reformulation of concepts, while encouraging the active participation of patients.

CONCLUSIONS

Our study, accomplished on a large cohort of undergraduate students in the context of ambulatory consultation, presents validity evidence to support the use of French-version SPSQ scores in screening for students experiencing difficulty with communication.

In the French-speaking medical world, the evaluation of professional skills has become a central element in the educational curriculum. This study is a significant step forward in French evaluation practices of medical students' interpersonal skills, especially because no validated questionnaire was available

before this study. In the endeavor to evaluate general skills in the ambulatory care environment, this questionnaire can be combined with a questionnaire evaluating clinical skills, such as the capacity to collect semiological data and to develop a clinical reasoning adapted to a patients' complaint.

In light of the fact that the tool has been designed to detect students experiencing difficulties with interpersonal skills, our results should be consolidated by two different kinds of study.

The first should aim at checking that the students with lower scores are truly those that experience the greatest difficulties. The second study should aim at exploring the fifth element, which supports validity according to Downing and the standards for educational and psychological testing,^{20,21} ie, the educational consequences of having implemented such a test. Because the SPSQ was initially designed for assessment of interpersonal skills in the United States, it will be essential to explore, eg, whether expectations of physician interpersonal skills with patients are different in France from in the United States.

Furthermore, in this perspective, a study evaluating the benefit of introducing coaching sessions for students recognized as experiencing difficulties with communication skills is currently ongoing within our medical faculty.

ACKNOWLEDGMENTS

The authors thank Mme Colette MOREL and Mme Alexandra REYNOLDS for the translation of the SPSQ.

REFERENCES

1. Issenberg SB, McGaghie WC, Petrusa ER, Lee Gordon D, Scalese RJ. Features and uses of high-fidelity medical simulations that lead to effective learning: a BEME systematic review. *Med Teach* 2005;27:10–28.
2. Gale AC, Gilbert J, Watson A. Life coaching to manage trainee underperformance. *Med Educ* 2014;48:539.
3. Campo AE, Williams V, Williams RB, Segundo MA, Lydston D, Weiss SM. Effects of LifeSkills training on medical students' performance in dealing with complex clinical cases. *Acad Psychiatry* 2008;32:188–193.
4. Sadr-Mohammadi R, Kalantari M, Molavi H. Efficacy of life skills training on subjective well-being of students: a report from rafsanjan, iran. *Iran J Psychiatry Behav Sci* 2014;8:63–67.
5. Pereira MA, Barbosa MA. Teaching strategies for coping with stress—the perceptions of medical students. *BMC Med Educ* 2013;13:50.
6. Wild K, Scholz M, Ropohl A, Bräuer L, Paulsen F, Burger PH. Strategies against burnout and anxiety in medical education—implementation and evaluation of a new course on relaxation techniques (Relacs) for medical students. *PLoS One* 2014;9:e114967.
7. Van Thiel J, Kraan HF, Van Der Vleuten CP. Reliability and feasibility of measuring medical interviewing skills: the revised Maastricht history-taking and advice checklist. *Med Educ* 1991;25:224–229.
8. Frankel RM, Stein T. Getting the most out of the clinical encounter: the four habits model. *J Med Pract Manage* 2001;16:184–191.
9. Kalet A, Pugnaire MP, Cole-Kelly K, et al. Teaching communication in clinical clerkships: models from the macy initiative in health communications. *Acad Med* 2004;79:511–520.
10. Kurtz S, Silverman J, Benson J, Draper J. Marrying content and process in clinical method teaching: enhancing the Calgary-Cambridge guides. *Acad Med* 2003;78:802–809.
11. Lang F, McCord R, Harvill L, Anderson DS. Communication assessment using the Common Ground Instrument: psychometric properties. *Fam Med* 2004;36:189–198.
12. Makoul G. Essential elements of communication in medical encounters: the Kalamazoo consensus statement. *Acad Med* 2001;76:390–393.

13. Makoul G. The SEGUE Framework for teaching and assessing communication skills. *Patient Educ Couns* 2001;45:23–34.
14. Lipner RS, Blank LL, Leas BF, Fortna GS. The value of patient and peer ratings in recertification. *Acad Med* 2002;77:S64–S66.
15. Stewart M, Brown JB, Donner A, et al. The impact of patient-centered care on outcomes. *J Fam Pract* 2000;49:796–804.
16. Novack DH, Dube C, Goldstein MG. Teaching medical interviewing: a basic course on interviewing and the physician-patient relationship. *Arch Intern Med* 1992;152:1814–1820.
17. Whelan GP. Educational Commission for Foreign Medical Graduates: clinical skills assessment prototype. *Med Teach* 1999;21:156–160.
18. Cohen DS, Colliver JA, Marcy MS, Fried ED, Swartz MH. Psychometric properties of a standardized-patient checklist and rating-scale form used to assess interpersonal and communication skills. *Acad Med* 1996;71: S87–S89.
19. Epstein RM, Dannefer EF, Nofziger AC, et al. Comprehensive assessment of professional competence: the Rochester experiment. *Teach Learn Med* 2004;16:186–196.
20. Downing SM. Validity: on meaningful interpretation of assessment data. *Med Educ* 2003;37:830–837.
21. American Educational Research Association, American Psychological Association, National Council on measurement in Education. *Standards for Educational and Psychological Testing*. Washington, DC: American Educational Research Association; 2014.
22. Schirmer JM, Mauksch L, Lang F, et al. Assessing communication competence: a review of current tools. *Fam Med* 2005;37:184–192.
23. Cleland JA, Abe K, Rethans JJ. The use of simulated patients in medical education: AMEE Guide No 42. *Med Teach* 2009;31:477–486.
24. Nickerson CAE. A note on “a concordance correlation coefficient to evaluate reproducibility.” *Biometrics* 1997;53:1503–1507.
25. Slocum-Gori SL, Zumbo BD. Assessing the unidimensionality of psychological scales: using multiple criteria from factor analysis. *Soc Indicators Res* 2011;102:443.
26. Crossley J, Davies H, Humphris G, Jolly B. Generalisability: a key to unlock professional assessment. *Med Educ* 2002;36:972–978.
27. Chessman AW, Blue AV, Gilbert GE, Carey M, Mainous AG 3rd. Assessing students' communication and interpersonal skills across evaluation settings. *Fam Med* 2003;35:643–648.
28. Wallace P. *Coaching Standardized Patients for the Use in the Assessment of Clinical Competence*. New York: Springer Publishing Company; 2007.

APPENDIX 1. The English Version of the SPSQ

	Rating Scale:				
	Poor	Fair	Good	Very Good	Excellent
Telling you everything; being truthful, up front, and frank; not keeping things from you that you should know.	1	2	3	4	5
Greeting you warmly, calling you by the name you prefer, being friendly, never crabby or rude.	1	2	3	4	5
Treating you like you're on the same level, never talking down to you or treating you like a child.	1	2	3	4	5
Letting you tell your story; listening carefully, asking thoughtful questions, not interrupting you while you're talking.	1	2	3	4	5
Showing interest in you as a person, not acting bored or ignoring what you have to say.	1	2	3	4	5
Discussing options with you, asking your opinion, offering choices and letting you help decide what you think before telling what to do.	1	2	3	4	5
Encouraging you to ask questions, answering them clearly, never avoiding your questions or lecturing you.	1	2	3	4	5
Explaining what you need to know about your problems—how and why they occurred and what to expect next.	1	2	3	4	5
Using words you can understand when explaining your problems and treatment, explaining any technical medical terms in plain language.	1	2	3	4	5
Understanding my feelings about my problems, appreciating the meaning of my problems to me.	1	2	3	4	5

APPENDIX 2. The French Translation of the SPSQ

	Médiocre	Passable	Bien	Très bien	Excellent
1. Le médecin me donne tous les éléments. Il est avenant et franc, il ne me cache pas d'éléments me concernant.					
2. Le médecin m'accueille de manière chaleureuse. Il s'adresse à moi d'une manière qui me convient. Le médecin est toujours courtois, jamais désagréable ou grossier.					
3. Le médecin se met à mon niveau. Il ne s'adresse pas à moi comme si j'étais inférieur ni me traite comme un enfant.					
4. Le médecin me laisse raconter mon histoire. Il est à mon écoute de manière attentive et me pose des questions attentionnées. Le médecin ne m'interrompt pas pendant que je parle.					
5. Le médecin montre de l'intérêt vis à vis de ma personne et n'agit pas comme s'il s'ennuie. Il n'ignore pas ce que j'ai à lui dire.					
6. Le médecin discute avec moi de différentes options, me demande mon avis, m'offre la possibilité de faire des choix et me laisse participer à la décision. Il me laisse exprimer ce que je pense avant de me dire quoi faire.					
7. Le médecin m'encourage à poser des questions et y répond clairement. Il ne les évite pas ni ne me fait de leçon.					
8. Le médecin m'explique ce que j'ai besoin de savoir concernant mes problèmes - pourquoi ils se sont produits et à quoi s'attendre.					
9. En utilisant des mots que je peux comprendre, le médecin m'explique mes problèmes et les traitements. Il explique les termes médicaux techniques dans un langage clair.					
10. Le médecin comprend mes sentiments au sujet de mes problèmes. Il reconnaît l'impact de mes problèmes.					