

# Cross-Cultural Adaptation of the Communication Assessment Tool for Use in a Simulated Clinical Setting

Citation for published version (APA):

Rocha, S. R., Romao, G. S., Setubal, M. S. V., Lajos, G. J., Luz, A. G., Collares, C. F., & Amaral, E. (2020). Cross-Cultural Adaptation of the Communication Assessment Tool for Use in a Simulated Clinical Setting. *Teaching and Learning in Medicine*, 32(3), 308-318.  
<https://doi.org/10.1080/10401334.2020.1717958>

## Document status and date:

Published: 26/05/2020

## DOI:

[10.1080/10401334.2020.1717958](https://doi.org/10.1080/10401334.2020.1717958)

## Document Version:

Publisher's PDF, also known as Version of record

## Document license:

Taverne

## Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

## General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

[www.umlib.nl/taverne-license](http://www.umlib.nl/taverne-license)

## Take down policy

If you believe that this document breaches copyright please contact us at:

[repository@maastrichtuniversity.nl](mailto:repository@maastrichtuniversity.nl)

providing details and we will investigate your claim.



## Cross-Cultural Adaptation of the Communication Assessment Tool for Use in a Simulated Clinical Setting

Sheyla Ribeiro Rocha, Gustavo Salata Romão, Maria Sílvia Velutini Setúbal, Giuliane Jesus Lajos, Adriana Gomes Luz, Carlos Fernando Collares & Eliana Amaral

To cite this article: Sheyla Ribeiro Rocha, Gustavo Salata Romão, Maria Sílvia Velutini Setúbal, Giuliane Jesus Lajos, Adriana Gomes Luz, Carlos Fernando Collares & Eliana Amaral (2020) Cross-Cultural Adaptation of the Communication Assessment Tool for Use in a Simulated Clinical Setting, *Teaching and Learning in Medicine*, 32:3, 308-318, DOI: [10.1080/10401334.2020.1717958](https://doi.org/10.1080/10401334.2020.1717958)

To link to this article: <https://doi.org/10.1080/10401334.2020.1717958>



Published online: 24 Feb 2020.



Submit your article to this journal [↗](#)



Article views: 210



View related articles [↗](#)



View Crossmark data [↗](#)

VALIDATION



## Cross-Cultural Adaptation of the Communication Assessment Tool for Use in a Simulated Clinical Setting

Sheyla Ribeiro Rocha<sup>a</sup> , Gustavo Salata Romão<sup>b</sup>, Maria Sílvia Velutini Setúbal<sup>c</sup>, Giuliane Jesus Lajos<sup>c</sup>, Adriana Gomes Luz<sup>c</sup>, Carlos Fernando Collares<sup>d</sup>, and Eliana Amaral<sup>c</sup>

<sup>a</sup>Department of Medicine, Federal University of São Carlos, São Carlos, Brazil; <sup>b</sup>Department of Medicine, University of Ribeirão Preto, Ribeirão Preto, Brazil; <sup>c</sup>Department of Obstetrics and Gynecology, University of Campinas, Campinas, Brazil; <sup>d</sup>Department of Educational Development and Research, Maastricht University, Maastricht, The Netherlands

### ABSTRACT

**Construct:** The Communication Assessment Tool (CAT) is a 14-item instrument developed in English to assess medical trainees' interpersonal communication skills from the patient's perspective in clinical settings. **Background:** Using validated instruments and simulated patients constitutes good practice in assessing doctor-patient communication. The CAT was designed for use in real practice, but has not yet been applied to assessing OB-GYN residents' delivery of bad news in Objective Structured Clinical Examination (OSCE) stations. This study aims to provide validity evidence for using the CAT to assess residents' interpersonal communication skills under difficult circumstances in a simulated clinical setting in Brazil. **Approach:** Cross-cultural adaptation comprised translation into Portuguese, synthesis of translations, and back-translation. Next, a committee of 10 external and independent experts rated the items for linguistic equivalence and relevance to the overall scale. Researchers used the expert ratings to produce a preliminary Brazilian-Portuguese version. This version was applied by four simulated patients to assess 28 OB-GYN residents completing two, 10-minute OSCE stations focused on delivering bad news. Item and scale content validity indices and internal-consistency reliability were calculated. Simulated patients were interviewed to clarify any doubt regarding the content and usability of the tool and their response process. **Findings:** Thirteen of the 14 items in the Brazilian-Portuguese version were considered "equivalent" by at least 70% of the experts. All items were considered relevant by 100% of the experts. The Item Content Validity Index ranged from .9 to 1, and the Scale Content Validity Index was .99. The instrument showed good reliability for both scenarios (Cronbach's alpha > .90). Simulated patients considered the CAT easy to understand and complete. **Conclusions:** This study provides validity evidence for using the Brazilian-Portuguese CAT in a simulated clinical environment to assess OB-GYN residents' delivery of bad news. Based on this study's findings, the OB-GYN Department organized an annual formative assessment for residents to improve their interpersonal communication skills. This version of the CAT may also be applicable to other specialties.

### KEYWORDS

cross-cultural adaptation; validation; communication skill; medical education; clinical assessment

## Introduction

Effective communication among health professionals, patients, and their families is recognized as a core aspect of quality health care.<sup>1</sup> Studies have shown that improving doctor-patient communication is directly related to better outcomes,<sup>2</sup> greater adherence to treatment,<sup>3</sup> greater satisfaction and reduction in malpractice complaints,<sup>4</sup> and improved patient safety, as well as reduction in errors.<sup>5</sup> Doctors who use communication skills (CS) effectively identify patients' problems more accurately, improve their own wellbeing, and have greater job satisfaction and less work stress.<sup>6</sup>

In light of these findings, global consensus recommends that teaching CS should be extended to all levels of medical training, including tasks and skills of the clinical interview.<sup>7,8</sup> Accordingly, since 2001, the Brazilian Curriculum Guidelines for Undergraduate Medical Education have included CS as an educational objective. To strengthen the acquisition of CS, repeated opportunities of assessment and feedback are recommended throughout vocational training.<sup>8</sup> Assessment of CS should include the patient's perspective, using Objective Structured Clinical Examination (OSCE) or workplace-based assessment.<sup>9</sup>

There are several tools available to assess CS. The Communication Assessment Tool (CAT) is a 15-item scale developed by Makoul et al. to assess doctors' interpersonal communication skills from the patient's perspective in clinical settings.<sup>10</sup> It is either self-administered or administered by an interviewer. Psychometric study has shown that the CAT has very high reliability (Cronbach's alpha = .96) and is unidimensional.<sup>10</sup> Differential item functioning analyses revealed that the CAT score was not affected by examinees' sociodemographic and clinical characteristics.<sup>10</sup> A 14-item version of the CAT also exists, which is tailored for medical students and residents.<sup>10</sup> This version focuses only on physician CS and omits the item related to respectful treatment of the patient by the doctor's staff. Since 2009, this version has been recommended by the Accreditation Council for Graduate Medical Education (ACGME) to assess medical residents in the USA.<sup>11</sup>

In the original study, Makoul et al.<sup>10</sup> reported data for the CAT from 38 doctors in different medical specialties. Following this, various residency programs such as family medicine,<sup>12,13</sup> internal medicine,<sup>14</sup> and general surgery<sup>15</sup> reported its use in clinical environments. Despite the widespread use of the CAT for workplace-based assessment, there are few examples of its use for assessing CS under difficult circumstances. Unpleasant announcements are an especially challenging aspect of doctor-patient interaction.<sup>16,17</sup> Acquiring the necessary skills for delivering bad news requires specific simulation-based training.<sup>16</sup> Nonetheless, the CAT has not been consistently applied in this context. To date, only one investigation has assessed the performance residents delivering bad news in a simulated scenario.<sup>18</sup>

Although the Brazilian Curriculum Guidelines for Undergraduate Medical Education recognizes CS as a core competency,<sup>19</sup> there is no validated tool to assess CS from the patient's perspective in simulated or clinical contexts in this country. Literature suggests that it is preferable to perform a cross-cultural adaptation of a validated instrument developed in another language than to create a new one, as it is relatively quicker and the results are comparable to other studies.<sup>20</sup> To face this challenge, the CAT was chosen based on its validation for use in measuring patient perceptions of CS performance of physicians-in-training.<sup>10</sup> Additionally, the CAT is written at a fourth-grade reading level and is therefore user-friendly for non-native English speakers.

There were two key objectives in this study. The first was to translate and adapt the 14-item version of

the CAT to the Brazilian context. The second was to collect validity evidence to use it as a formative assessment of OB-GYN residents delivering bad news in simulated environments. Four sources of validity evidence were investigated: (1) evidence regarding test content provided by experts' judgment regarding the relationship between instrument items and the CS construct; (2) evidence regarding internal structure by calculating the Cronbach's alpha coefficient, an index of internal-consistency reliability; (3) evidence regarding response process by interviewing simulated patients after the use of the adapted scale to assess residents' performance; (4) evidence regarding consequences after discussing the results and the repercussions of this study.

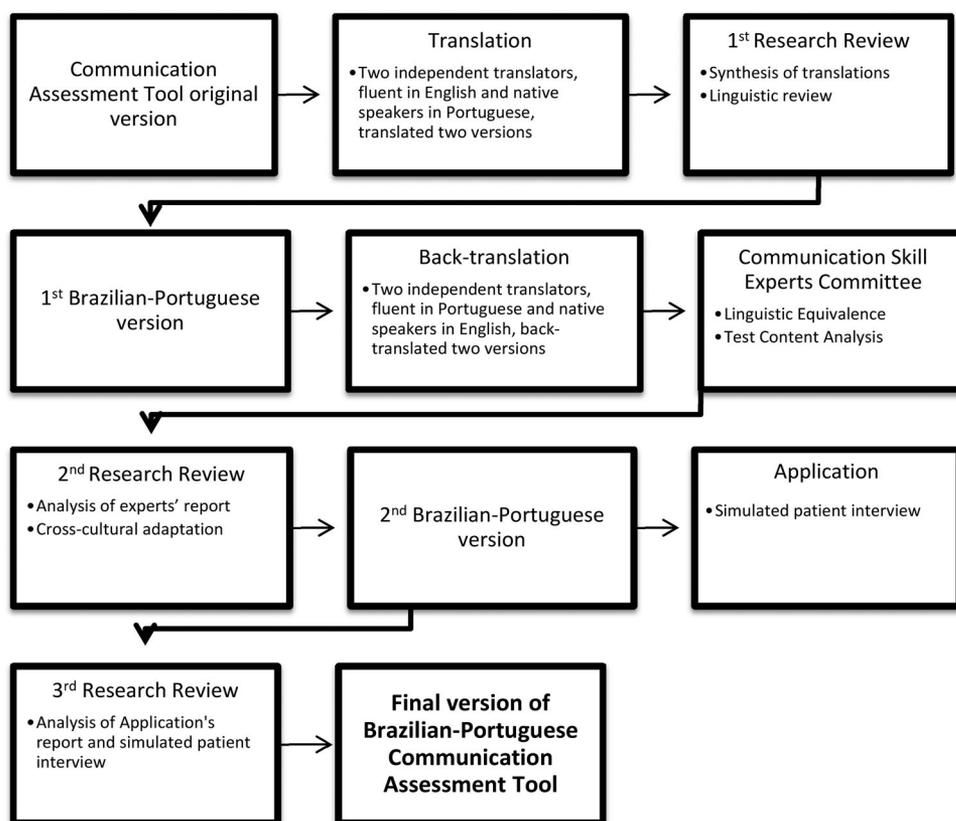
## Method

### *Translation and cross-cultural adaptation of the CAT*

A variety of protocols are used to perform cross-cultural adaptation of instruments, each with advantages and drawbacks. In the present study, we adopted a multistep protocol based on previous studies and well-established guidelines for the cross-cultural adaptation of instruments<sup>20-22</sup> (see Figure 1). Prior to implementing the protocol, we obtained the developer's permission to use the instrument. Then, our research proposal was presented to the Institutional Human Ethical Committee at the University of Campinas (CAAE 56935716.5.0000.5404).

This research was conducted between 2016 and 2018. The protocol involved the following steps: (1) translation of the original instrument to the target language; (2) linguistic review and synthesis of the translations to produce the first version of the instrument; (3) back-translation; (4) analyses of the linguistic equivalence and test content by expert committee; (5) review of the analysis of experts' report by the researchers to produce the second version of instrument; (6) application of the adapted version; (7) cognitive interview with the simulated patient; and (8) production of the final version of the adapted instrument. A detailed description of this multistep protocol follows.

First, two-independent official translators who were fluent in English and were native Portuguese speakers translated two versions (T1 and T2) of the CAT. Next, a research group comprising four medical education experts and one CS expert met virtually and face-to-face to compare these two translations and develop a synthesized Brazilian-Portuguese version



**Figure 1.** Step sequence for the translation and cross-cultural adaptation process of the Brazilian-Portuguese Communication Assessment Tool.

**Table 1.** Demographic characteristics of Communication Skill Experts Committee.

Feature	N	%	Mean (years)
Sex			
Female	6	60	
Male	4	40	
Degree			
Doctorate	6	60	
Master	4	40	
Teaching in Medical Education			25
4–20 years	3	30	
21–36 years	7	70	
Teaching Communication Skills			14
2–20 years	7	70	
21–36 years	2	20	
Experience in Scale Adaptation			
Yes	4	40	
Brazilian Region			
Northeast	1	10	
South	2	20	
Southeast	7	70	

(T12). Inappropriate words and expressions were modified by consensus, and the entire process was carefully documented. The synthesis version was back translated by two-independent professionals who were fluent in Portuguese and were native English speakers (BT1 and BT2). None of the translators had access to the original instrument.

Throughout this process, 10 external and independent experts were invited to form a committee, and written informed consent was provided by all experts. They were selected based on their experience in teaching CS in various schools and regions in Brazil (see Table 1). The Communication Skill Experts (CSEs) assessed the equivalence between the synthesized version and the original instrument in four different areas: semantic, idiomatic, experiential, and conceptual.<sup>20</sup> A rating scale was used as follows: 1 (nonequivalent), 0 (indecisive), and +1 (equivalent). Each item was considered equivalent as long as 7 out of 10 CSEs regarded it as so.<sup>23</sup>

In this step, CSEs could propose modifications and new translations that better suited the CAT Brazilian-Portuguese version. They received guidance on the objective of the research, the concepts regarding cross-cultural adaptation, the description of the instrument, as well as the versions (the original CAT, T12, BT1, and BT2). They also received a compendium and a rationale written by the research group that created the synthesized version. This report provided detailed information such as the lexical meaning and word adequacy in the original and target language. The CSEs were provided assistance if doubts arose during the process.

The research group assessed the complete written documents provided by the CSEs and made critical decisions concerning the Brazilian-Portuguese version of the CAT. Each item was mapped by the researchers; regular meetings occurred to discuss the data and reach a consensus regarding the best adaptation.

### **Evidence of content validity**

A quantitative approach was used to examine content validity: the Content Validity Index (CVI).<sup>24,25</sup> The CSEs rated each CAT item on their construct relevance. The following rating scale was used as recommended by Polit et al.<sup>24</sup> 1 = not relevant; 2 = somewhat relevant; 3 = quite relevant; 4 = highly relevant. Then, for each item, an Item-level Content Validity Index (I-CVI) was computed by dividing the number of experts who had rated the item a 3 or 4 (relevant) by the total number of experts. Specifically, the I-CVI expresses the proportion of agreement on the relevance of each item among the CSEs. The CSEs also evaluated the CAT items for clarity using yes/no questions. Each item was considered clear if 7 out of 10 CSEs regarded it as so.<sup>23</sup>

### **Instrument application**

#### **Setting and population study**

The Center for Women's Health, which manages the OB-GYN Residency Program, consists of a public, quaternary, and tertiary University Hospital and an outpatient clinic at the University of Campinas. The 3-year OB-GYN Residency Program offers 11 resident positions per year. The Obstetric Emergencies Course is mandatory for residents in the first-year of their training and optional for their second and third years. It consists of an annual, one-day workshop, where simulation techniques are used to teach residents to deal with the typical clinical conditions of obstetric emergencies. This CS training was part of a pilot project in which residents volunteered to participate in doctoral research carried out between 2014 and 2017.<sup>26</sup> The training consisted of delivering bad news using SPIKES, a protocol for disclosing unfavorable information.<sup>27</sup> Until 2019, CS training had not been part of the OB-GYN residency curriculum in Brazil.<sup>26</sup> For this study, the syllabus of this course included breaking bad news during antenatal care for the first time.

A total of 33, first-year through third-year residents were eligible to participate in the training. Out of these, 28 residents were present in the workshop. All

28 residents who attended the course were invited and informed about the study and the formative feedback they would receive. All of them agreed and provided written informed consent. After that, they answered a questionnaire regarding their sex, postgraduate year, and previous CS training. They were assured that their CS assessment would not be considered in progress decision making. Therefore, no weight was given to this OSCE for performance evaluation purposes.

### **Designing the communication scenarios**

The scenarios used for this study were designed with the curriculum of the OB-GYN Brazilian Residency Program in mind,<sup>28</sup> as well as the Protocols of the Brazilian Ministry of Health concerning obstetric care to prevent maternal complications.<sup>29</sup> Both OSCE stations were designed by the researchers and professors in charge of the Obstetric Emergency Course.

The task at Station 1 was to inform a female patient of a serologic diagnosis of syphilis during her first prenatal visit in an outpatient clinic. The task of Station 2 consisted of informing a pregnant woman in the maternity hospital of an unexpected diagnosis of preeclampsia with high risk of fetal loss. The objective of the residents in both situations was to communicate the diagnosis to the patient and to discuss the care plan and the next steps. Because the purpose of this research was to adapt CAT to be used in the Brazilian context, an OSCE with two different stations was considered sufficient.<sup>20</sup>

### **Standardized patients' selection, training, and interviews**

Four simulated patients (SPs) were recruited from a group of well-trained SPs from the Medical School at the University of Campinas. They were recruited based on their previous experience in acting as SPs in undergraduate and residency assessment tests. They received a 3-hour training workshop on using the CAT and performing their role. First, they read the items, and their understanding was verified. Then, they received guidance on completing the CAT. After that, the researchers and the SPs discussed what to expect regarding resident behavior, as well as the role of the SP in the scenario.

After the OSCE, the SPs were individually interviewed by one of the researchers. The interviewer used two cognitive interviewing approaches: verbal probing and think-aloud techniques. Each interview took between 15 and 20 minutes. The interview aimed to clarify any doubt regarding the content of the tool,<sup>20</sup> the usability in a simulated scenario,<sup>22</sup> and the

SP response process.<sup>30</sup> Unlike the other participants, the SPs received payment for this activity. They also provided written, informed consent.

### Scoring process

The SPs rated residents' CS using the Brazilian-Portuguese CAT immediately after their performance. Each item was scored on a five-point scale (1="poor," 2="fair," 3="good," 4="very good," and 5="excellent"). The overall score was obtained in two ways: the percentage of items rated as "Excellent" and the sum of the item ratings (ranging from 14 to 70). The developers suggested that scoring based on the proportion of items rated "Excellent" by patients was more useful than summarizing scores via means, which are highly biased.<sup>10</sup> However, in other studies involving residents, both score processes were reported.<sup>12,13,15,18</sup> In this study, data were analyzed using both methods.

### Running the OSCE

All 28 residents participated in two, 10-minute OSCE stations focused on CS under difficult circumstances. At the end of each station, SPs completed the CAT for one minute. The OSCE ran in seven rotations with two simultaneous circuits comprising the same two stations. At the end of the OSCE, residents and SPs debriefed together for 40 minutes, discussing the CAT Brazilian-Portuguese items and SP perceptions. The debriefing was facilitated by two researchers.

### Data analysis

Data analysis involved diverse procedures according to the multistep protocol used in this research as follows: (1) estimation of the linguistic equivalence and clarity during the translation and cross-cultural adaptation; (2) estimation of the content validity indices; (3) estimation of the reliability index; (4) analysis of residents' performance data; and (5) analysis of the SP cognitive interviews.

### Linguistic equivalence and clarity

Linguistic equivalence for each item was obtained by calculating the proportion of CSEs who rated it as "Equivalent." An item was considered equivalent as long as 7 out of 10 CSEs regarded it as so.<sup>23</sup> Item clarity was obtained by the proportion of CSE's agreement with a yes/no statement that the item was clear. Each item was considered clear as long as 7 out of 10 CSE regarded it as so.<sup>23</sup>

### Content validity index

The Content validity index (CVI) was applied for each item (I-CVI) and for the entire scale (S-CVI).<sup>24</sup> The I-CVI was obtained as the proportion of CSEs who rated the item as relevant (a 3 or 4 on the relevance rating scale), divided by the total number of CSEs.<sup>24</sup> The S-CVI was obtained by computing the I-CVI for each item on the scale, and then calculating the average I-CVI across items.<sup>24</sup> The CVI ranges from 0 to 1. Because there were 10 CSEs, the Brazilian-Portuguese CAT was considered as having excellent content validity if items showed an I-CVI of .78 or greater, and a S-CVI of .90 or greater.<sup>24,31</sup>

### Reliability index

Internal consistency was estimated using Cronbach's alpha coefficients. A value  $\geq 0.7$  was considered good reliability.<sup>32</sup>

### Residents' performance data

Residents' performance scores were obtained in two ways: the percentage of items rated as "Excellent" and the sum of the item ratings. The resident score was stratified based on gender, previous CS training, and postgraduate year. Nonparametric analysis was performed using Mann-Whitney test for gender and previous CS training, and Kruskal-Wallis test for postgraduate year. For overall CAT ratings, we also examined the mean score and percentage of "Excellent" scores for each item. Statistical significance was defined as a  $p$  value less than .05. Statistical analysis was performed using GraphPad Prism 8.

### Cognitive interviews with SPs

SP interviews were audio-recorded following the interviewees' consent. Data were analyzed using a confirmatory thematic analysis approach.<sup>33</sup> The interviews were coded by a single researcher to reveal issues concerning (a) understanding items, (b) the scale context, and (c) retrieving and integrating the information used to score the resident's performance.<sup>34</sup> Afterward, the research group reviewed and discussed the codes to mitigate any potential bias.

## Results

### Linguistic equivalence and clarity

Ten CSEs rated the translation and cross-cultural adaptation of each of the 14 Brazilian-Portuguese CAT items using a three-point rating scale. Nine items were considered "Equivalent" by at least 70% of the CSEs; four items (2, 6, 9, and 10) were considered

**Table 2.** Communication Assessment Tool Scores.

CAT item	Mean (SD)	Median	Range	% of items rated as "Excellent"
1. Greeted me in a way that made me feel comfortable	2.89 (1.15)	3	1–5	8.9
2. Treated me with respect	2.96 (0.92)	3	1–5	5.5
3. Showed interest in my ideas about my health	2.45 (1.04)	2	1–5	1.8
4. Understood my main health concerns	2.32 (0.99)	2	1–5	1.8
5. Paid attention to me (looked at me, listened carefully)	2.96 (1.06)	3	1–5	7.1
6. Let me talk without interruptions	2.79 (0.99)	3	1–5	1.8
7. Gave me as much information as I wanted	2.68 (0.96)	3	1–5	3.6
8. Talked in terms I could understand	2.68 (1.10)	3	1–5	5.4
9. Checked to be sure I understood everything	2.50 (1.06)	2	1–5	1.8
10. Encouraged me to ask questions	2.14 (1.05)	2	1–4	0
11. Involved me in decisions as much as I wanted	2.05 (0.87)	2	1–4	0
12. Discussed next steps, including any follow-up plans	2.71 (0.97)	3	1–5	7.1
13. Showed care and concern	2.61 (1.00)	2	1–5	7.1
14. Spent the right amount of time with me	2.58 (1.06)	2	1–5	3.8
Average score	36.1 (8.1)	36.3	18–56	4

“equivalent” by 100% of the CSEs. Only item 3 (*Showed interest in my ideas about my health*) obtained lower agreement (60%) regarding cultural equivalence due to the meaning of the word “ideas” in the Brazilian context. This item was rewritten according to the CSEs’ suggestions, and the word “idea” was replaced with “opinions.” All items were considered clearly written by 100% of CSE. No further adaptation was necessary.

### Content validity indices

All 14 items were considered associated with the construct, and “relevant” by 100% of the CSEs. The I-CVI ranged from .90 to 1; the S-CVI was .99. Based on this data, all items were maintained in the Brazilian-Portuguese CAT version. This version was used to assess the OB-GYN residents in a simulated environment.

### Reliability

The reliability of the 14-item CAT was .95. The Cronbach’s alpha coefficient was also calculated separately at each station to confirm this high value. Data analysis revealed that coefficients were .93 and .95 for Station 1 and 2, respectively. These results are in accordance with those observed in previous research findings.<sup>10,15,35</sup>

### Residents’ performance data

The Brazilian-Portuguese CAT was completed by 4 SPs for 28 OB-GYN residents (11 PGY-1, 9 PGY-2, and 8 PGY-3; 78.5% female) during two OSCE stations. Two CAT forms were completed per resident by two SPs. A total of 71.4% of participants self-reported previous CS training during undergraduate

studies (8/28), OB-GYN residency (9/28), or both (3/28). Most of the residents (92.8%) had already participated in at least one OSCE assessment of clinical skills.

Residents’ scores in the CAT, averaged across scenarios, ranged from 18 to 56 ( $M=36.1$ ,  $Mdn=36.3$ ,  $SD=8.1$ ). The average proportion of “Excellent” scores ranged from 0% to 32%; only 29% (8/28) of the residents received at least one item rated as “Excellent.” Among those, the proportion of “Excellent” ratings ranged from 4% to 32% of the items. This performance was much lower than expected.

No significant effect of trainee gender was found (female:  $Mdn=36$ ; male:  $Mdn=39.25$ ),  $U=54$ ,  $p=.520$ ,  $r=-0.24$ . Contrary to expectations, further analyses revealed no significant effect of previous CS training (for trained and not trained, respectively  $Mdn=36.5$  and  $Mdn=35.0$ ),  $U=59$ ,  $p=.297$ ,  $r=.24$ . No significant effect of postgraduate year was found (PGY1:  $Mdn=34.5$ ; PGY2:  $Mdn=36.5$ ; PGY3:  $Mdn=38$ ),  $H(2, 27)=.95$ ,  $p=.623$ .

Table 2 presents the mean scores, median, and the percentage of “Excellent” ratings for CAT items. The data reveals that the highest scored items were “Treated me with respect” ( $M=2.96$ ,  $Mdn=3$ ,  $SD=0.92$ ), “Paid attention to me (looked at me, listened carefully)” ( $M=2.96$ ,  $Mdn=3$ ,  $SD=1.06$ ), and “Greeted me in a way that made me feel comfortable” ( $M=2.89$ ,  $Mdn=3$ ,  $SD=1.15$ ). The lowest scored items were “Involved me in decisions as much as I wanted” ( $M=2.05$ ,  $Mdn=2$ ,  $SD=0.87$ ) and “Encouraged me to ask questions” ( $M=2.14$ ,  $Mdn=3$ ,  $SD=1.05$ ).

### Cognitive interviews with SPs

SPs comprised four females with Bachelor of Arts degrees. All of them had previous experience (6.7 years in average) in acting as a patient in undergraduate and

residency assessment tests. They had already received feedback/debriefing training in interactions where bad news was delivered (SPIKES). They were also familiar with research projects in CS in medical education.

Following the application of the Brazilian-Portuguese CAT, three SPs were individually interviewed. First, they read each item aloud and their understanding of the item was verified. They considered it easy to understand and fill out. After that, none expressed any doubt regarding the items' content and the instrument's form. Considering the CAT's usability in a simulated environment, a common view among the interviewees was that the CAT was user friendly.

I found the scale very detailed and helpful. It is logical and, then, we pay more attention to the details, when you know what to observe. (SP3)

However, when asked about the length of time to complete the instrument, one SP argued that it would have been more "comfortable" if she had had more time to assess the resident's performance.

SPs agreed that the behaviors described by Items 11 ("*Involved me in decisions as much as I wanted*") and 12 ("*Discussed next steps, including follow-up plans*") were considered the most difficult ones to assess. There was a sense among the SPs that these items were broader and more subjective than the others.

Items 11 and 12 raised some doubts. Because, many times, the next steps were related to decisions she was taking.... Many times, she [the resident] didn't discuss with me. She only let me know the decisions. (SP1)

It was very difficult for me to know if she had been successful or not [referring to Items 11 and 12] because I didn't know what she should discuss, exactly. This is a little difficult to assess because sometimes she shared a possibility with me. Sometimes she asked questions. Sometimes she made a recommendation. Sometimes she was more assertive. Then, I didn't know to what extent she was discussing (the next steps) with me or just letting me know about the situation. (SP2)

Some SP comments regarding their reasoning for rating the residents' performance could be linked with their training in completing the CAT. They reported using a pattern to assess the resident's performance. Some of them expressed their concerns regarding fairness during the assessment.

I thought about how not to be unfair but followed a pattern [referring to the training]. There are many ways for a professional to show these behaviors. (SP3)

As the SPs rated the residents at the end of the station, some reported using their memories of what had

happened during the interaction in the rating process. One SP reported having an overall perception at the end of the appointment based on her feelings; this perception guided the assessment of each item. Somehow, all of them used their emotions and feelings to assess residents' performance. They justified this based on the idea that the behaviors in the scale were more than a simple task.

It [the resident's assessment] was based on the way I felt. The way this patient would feel. Then, I thought about the relationship itself and not about the tasks she performed. Because, sometimes, although the person greeted me, she didn't make me feel comfortable. It was indeed subtler... my perception. (SP1)

I (rated) based on the memory of what had just happened and the sensation regarding the appointment as she left the room. As an actress, this is really intense. By the time the appointment was over, the general sensation also impacted the assessment of the individual questions... (SP2)

No adaptation was required after the SP interviews. The final version of the Brazilian-Portuguese CAT is presented in the appendix.

## Discussion

The use of validated instruments and the inclusion of simulated patients' perceptions are recommended as good practice to assess doctor-patient communication.<sup>36,37</sup> With this in mind, we chose to translate and adapt the CAT to assess CS under difficult circumstances in simulated environments. According to our results, the CAT was an appropriate choice for the formative assessment of Brazilian medical residents in this scenario.

The process of cross-cultural adaptation was conducted according to the standards recommended by international literature.<sup>20-22</sup> In addition to linguistic equivalence and clarity, content validity, reliability, and utility/feasibility in a simulated environment also were verified. All items in the original instrument were maintained in the Brazilian-Portuguese CAT version, considering the high degree of endorsement by the CSEs.<sup>23,31</sup> Only one item was adapted to the Brazilian context, in accordance with CSEs' reports.

There is no consensus on the ideal strategy for the transcultural adaptation of instruments. Methodological choice has been claimed a matter of preference and logistics.<sup>22,38</sup> Most protocols recommend checking the content validity while verifying cross-cultural adaptation. Nevertheless, they rarely define the steps to achieve this objective.<sup>38</sup> Our protocol sheds some light on this matter, since the cross-cultural equivalence,

clarity, and item relevance were verified by 10 CSEs. For this reason, the CSEs were selected carefully; all were experienced teachers from different Brazilian regions. This ensured that the Brazilian-Portuguese CAT version contained only adequate and relevant items.<sup>23</sup>

Use of the CAT for workplace-based assessment has been validated.<sup>10,12,13,15,39</sup> However, there is little experience concerning its use in a simulated environment assessing CS under difficult circumstances. Based on the SPs' perspective, the Brazilian-Portuguese CAT is a user-friendly tool for OSCEs.

The reliability of the CAT was high in both stations for this group of OB-GYN residents (Cronbach's alpha coefficient=.93 and .95). These results mirror previous studies using CAT in real<sup>10,35,39</sup> and simulated environments.<sup>18</sup> Owing to the small sample, an alternative measurement of reliability was not supposed to be performed. The Cronbach's alpha coefficient was reported here to support the assumption that the 14-item CAT version is reliable for this specific group of residents in the Brazilian context and to draw a comparison with other studies.<sup>32</sup>

One interesting finding involved the highest and lowest rated items. The two highest rated items were "Treated me with respect" ( $M=2.96$ ) and "Paid attention to me (looked at me, listened carefully)" ( $M=2.96$ ). These are the same items reported by Makoul et al.,<sup>10</sup> Myerholtz et al.,<sup>12,13</sup> and Stausmire et al.<sup>15</sup> The two lowest rated items were "Involved me in decisions as much as I wanted" ( $M=2.05$ ) and "Encouraged me to ask questions" ( $M=2.14$ ). These results are consistent with those reported by Makoul et al.,<sup>10</sup> Myerholtz et al.,<sup>12,13</sup> Stausmire et al.,<sup>15</sup> Trickey et al.,<sup>18</sup> and Ferranti et al.<sup>39</sup> Our data indicates that regardless of the cultural background or the type of CS training received, the same CAT items repeatedly appear in the highest and the lowest scores. These results reveal areas for improvement during residents' training. It involves training the residents to give patients voice to ask questions and to be involved in decisions regarding their own lives, which embodies the principles of patient-centered care.<sup>40</sup>

Contrary to expectations, the performance of the residents was very low. The average proportion of "Excellent" scores was 4%, while previous studies reported an average proportion of "Excellent" scores at 70%.<sup>10,12,15,39</sup> Although these results differ from some published studies in daily clinical practice, they corroborate those of Trickey et al.<sup>18</sup> In that study, general surgery residents were assessed while delivering bad news in a simulated environment.<sup>18</sup> The average proportion of "Excellent" ratings was 5%. In

addition, performance data showed no impact of previous CS training or postgraduate year.

A possible explanation for these results may be the lack of adequate CS training during the OB-GYN Residency Program. In our sample, almost 30% (8/28) of the residents self-reported that they had never attended formal CS training. Several authors have discussed the challenges and strategies for supporting medical students in becoming effective communicators.<sup>40-42</sup> There is a consensus that the best way to promote acquisition of relevant CS is to establish an effective training program and a coherent assessment approach in undergraduate and postgraduate teaching.<sup>1,7,41,43</sup>

Another plausible explanation for lower scores is that the task was performed in an OB-GYN setting. In this specialty, the communication of difficult news is particularly challenging due to the high expectations involved in having a child.<sup>44,45</sup> Setubal et al. highlight the need for training in delivering bad news during the OB-GYN Residency in Brazil.<sup>26</sup> Our results corroborate these findings since only 43% (12/28) of the residents self-reported having attended a SPIKES training during their residency program. It is important to note that, as recent as 2019, the Brazilian Competency Framework for OB-GYN Residency presented CS as an essential competency, especially in delivering bad news.<sup>46</sup> Before that, no specific recommendation on CS training in an OB-GYN residency program existed.<sup>26</sup>

Another explanation for the lower ratings involves the training of the SPs, as reported by Trickey et al.<sup>18</sup> In our study, according to the interviews with the SPs, the behaviors described by the items "Involved me in decisions as much as I wanted" and "Discussed next steps, including follow-up plans" were considered the most difficult ones to assess because these items are broader and more subjective. Somehow, all of the SPs addressed their emotions and feelings to assess the residents' performance. They justified this by noting that the behaviors assessed by the scale were more than simple tasks. That said, study has revealed that CS assessment by SPs can be highly correlated with the assessment of medical students in real environments.<sup>37</sup>

Because communication is culturally mediated and involves a combination of values, knowledge, and behavioral skills, some issues could potentially be addressed. According to Morrow et al., Hofstede's framework and cultural dimensions may be helpful in understanding cultural norms and values in the doctor-patient relationship and communication.<sup>42</sup> The power distance dimension (related to the extent to which power is distributed in the society) may impact

the type and extent of communication doctors have with patients and, consequently, how patients perceive the communication.<sup>42</sup> With this in mind, we propose further investigation into the role of a trained SP and consider her perceptions and reasoning to assess the physicians' performance considering the Hofstede's framework.

The present study has several limitations. First, it is the initial evaluation of validity evidence for a Brazilian-Portuguese version of the CAT. Independent replication of these results on a larger scale is required in other institutions and contexts. Second, we used a small sample and only two OSCE stations due to restrictions on time and financial resources. For the same reason, we designed the educational intervention with only one SP at each station. Therefore, further evidence found by comparing inter-rater agreement was not provided. Third, we did not have another validated scale for CS assessment in Brazilian medical education at the time of this research. Consequently, we could not design a convergent study. Finally, the researcher responsible for interviewing the SPs and coding these data was also involved in the SPs training. To address this potential for bias, the research group reviewed and discussed the codes.

The process of validation aimed to collect relevant evidence to provide a scientific basis to support the decision on resident's performance. This process should contain multiple sources of evidence that allow some inferences on the validity argument.<sup>47</sup> The present study built additional validity evidence for the use of the CAT in a simulated environment to assess residents' CS based on several sources. First, the CSEs evaluated the item relevance and the high experts' agreement provided validity evidence based on test content. Second, calculating the Cronbach's alpha coefficient showed high reliability for this sample of residents. Third, the SPs' cognitive interviews offered validity evidence based on their response processes when assessing the residents' performance. Finally, following our findings, the OB-GYN Department organized an annual formative assessment for residents to improve their skills in dealing with doctor-patient relationship.

Taken together, these results suggest that the Brazilian-Portuguese CAT is a valid and reliable option to assess medical residents' CS in Brazil. The CAT proved to be a simple and feasible instrument when used in a simulated environment by trained SPs. A natural progression of this work is to investigate the use of the CAT on larger samples and in different scenarios. The findings of this research highlighted the need for efforts to ensure that all

Brazilian postgraduate teaching offers CS training and provides opportunities to practice these skills under controlled and safe conditions.

### Declaration of interest statement

The authors have no conflicts of interest to declare.

### Funding

This study was supported by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES).

### ORCID

Sheyla Ribeiro Rocha  <http://orcid.org/0000-0003-1926-4439>

### References

1. Laidlaw A, Hart J. Communication skills: an essential component of medical curricula. Part I: Assessment of clinical communication: AMEE Guide No. 51. *Med Teach.* 2011;33(1):6–8. doi:10.3109/0142159X.2011.531170.
2. Stewart MA. Effective physician-patient communication and health outcomes: a review. *CMAJ.* 1995; 152(9):1423–1433.
3. Haskard Zolnieriek KB, DiMatteo MR. Physician communication and patient adherence to treatment: a meta-analysis. *Med Care.* 2009;47(8):826–834. doi:10.1097/MLR.0b013e31819a5acc.
4. Levinson W, Roter DL, Mullooly JP, et al. Physician-patient communication: The relationship with malpractice claims among primary care physicians and surgeons. *JAMA.* 1997;277(7):553–559. doi:10.1001/jama.1997.03540310051034.
5. Sandars J, Bax N, Mayer D, et al. Educating undergraduate medical students about patient safety: Priority areas for curriculum development. *Med Teach.* 2007; 29(1):60–61. doi:10.1080/01421590601087546.
6. Maguire P, Pitceathly C. Key communication skills and how to acquire them. *Br Med J.* 2002;325(7366): 697–700. doi:10.1136/bmj.325.7366.697.
7. Kiessling C, Dieterich A, Fabry G, et al. Ausbildung CCaScotAfMEGfM and others. Communication and social competencies in medical education in German-speaking countries: the Basel consensus statement. Results of a Delphi survey. *Patient Educ Couns.* 2010; 81(2):259–266. doi:10.1016/j.pec.2010.01.017.
8. von Fragstein M, on behalf of the UK Council for Clinical Communication Skills Teaching in Undergraduate Medical Education, Silverman J, Cushing A, Quilligan S, Salisbury H, Wiskin C. UK consensus statement on the content of communication curricula in undergraduate medical education. *Med Educ.* 2008;42(11):1100–1107. doi:10.1111/j.1365-2923.2008.03137.x.

9. Kiessling C, Tsimtsiou Z, Essers G, et al. General principles to consider when designing a clinical communication assessment program. *Patient Educ Couns*. 2017;100(9):1762–1768. doi:10.1016/j.pec.2017.03.027.
10. Makoul G, Krupat E, Chang CH. Measuring patient views of physician communication skills: development and testing of the Communication Assessment Tool. *Patient Educ Couns*. 2007;67(3):333–342. doi:10.1016/j.pec.2007.05.005.
11. Swing SR, Clyman SG, Holmboe ES, et al. Advancing resident assessment in graduate medical education. *J Grad Med Educ*. 2009;1(2):278–286. doi:10.4300/JGME-D-09-00010.1.
12. Myerholtz L. Assessing family medicine residents' communication skills from the patient's perspective: evaluating the communication assessment tool. *J Grad Med Educ*. 2014;6(3):495–500. doi:10.4300/JGME-D-13-00347.1.
13. Myerholtz L, Simons L, Felix S, et al. Using the communication assessment tool in family medicine residency programs. *Fam Med*. 2010;42(8):567–573.
14. Wayne DB, Cohen E, Makoul G, et al. The impact of judge selection on standard setting for a patient survey of physician communication skills. *Acad Med*. 2008;83(Supplement):S17–S20. doi:10.1097/ACM.0b013e318183e7bd.
15. Stausmire JM, Cashen CP, Myerholtz L, et al. Measuring general surgery residents' communication skills from the patient's perspective using the Communication Assessment Tool (CAT). *J Surg Educ*. 2015;72(1):108–116. doi:10.1016/j.jsurg.2014.06.021.
16. Johnston FM, Beckman M. Navigating difficult conversations. *J Surg Oncol*. 2019;120(1):23–29. doi:10.1002/jso.25472.
17. Farber NJ, Urban SY, Collier VU, et al. The good news about giving bad news to patients. *J Gen Intern Med*. 2002;17(12):914–922. doi:10.1046/j.1525-1497.2002.20420.x.
18. Trickey AW, Newcomb AB, Porrey M, et al. Assessment of Surgery Residents' Interpersonal Communication Skills: Validation Evidence for the Communication Assessment Tool in a Simulation Environment. *J Surg Educ*. 2016;73(6):e19–e27.
19. Diretrizes Curriculares Nacionais para Cursos de Graduação em Medicina [Medical Courses' Curriculum Guidelines] [Internet]. Brasil: Ministério da Educação, Conselho Nacional de Educação, Câmara de Educação Superior. 2001 [cited 2019. Jan 20]. Available from: <http://portal.mec.gov.br/cne/arquivos/pdf/CES04.pdf> Portuguese.
20. Beaton DE, Bombardier C, Guillemin F, et al. Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine (Phila Pa 1976)*. 2000;25(24):3186–3191. doi:10.1097/00007632-200012150-00014.
21. Gjersing L, Caplehorn JR, Clausen T. Cross-cultural adaptation of research instruments: language, setting, time and statistical considerations. *BMC Med Res Methodol*. 2010;10(1):13. doi:10.1186/1471-2288-10-13.
22. Reichenheim ME, Moraes CL. [Operationalizing the cross-cultural adaptation of epidemiological measurement instruments. *Rev Saúde Pública*. 2007;41(4):665–673.
23. Sireci S, Faulkner-Bond M. Validity evidence based on test content. *Psicothema*. 2014;26(1):100–107. doi:10.7334/psicothema2013.256.
24. Polit DF, Beck CT, Owen SV. Is the CVI an acceptable indicator of content validity? Appraisal and recommendations. *Res Nurs Health*. 2007;30(4):459–467. doi:10.1002/nur.20199.
25. Lynn MR. Determination and quantification of content validity. *Nurs Res*. 1986;35(6):382–385.
26. Setubal MSV, Gonçalves AV, Rocha SR, et al. Breaking Bad News Training Program Based on Video Reviews and SPIKES Strategy: What do Perinatology Residents Think about It? *Revista Brasileira de Ginecologia e Obstet*. 2017;39(10):552–559. doi:10.1055/s-0037-1604490.
27. Baile WF, Buckman R, Lenzi R, et al. SPIKES-A six-step protocol for delivering bad news: application to the patient with cancer. *Oncologist*. 2000;5(4):302–311. doi:10.1634/theoncologist.5-4-302.
28. Requisitos Mínimos dos Programas de Residência Médica [Medical Residency' Curriculum Guidelines] [Internet]. Brazil: Ministério da Educação, Conselho Nacional de Educação, Câmara de Educação Superior. 2006 [cited 2019. Jan 30]. Available from: [http://portal.mec.gov.br/dmdocuments/resolucao02\\_2006.pdf](http://portal.mec.gov.br/dmdocuments/resolucao02_2006.pdf) Portuguese.
29. Pacagnella RC, Nakamura-Pereira M, Gomes-Sponholz F, et al. Moraes Filho OBd. Maternal mortality in Brazil: proposals and strategies for its reduction. *Rev Bras Ginecol Obstet*. 2018;40(09):501–506. doi:10.1055/s-0038-1672181.
30. Padilla JL, Benitez I. Validity evidence based on response processes. *Psicothema*. 2014;26(1):136–144. doi:10.7334/psicothema2013.259.
31. Polit DF, Beck CT. The content validity index: are you sure you know what's being reported? Critique and recommendations. *Res Nurs Health*. 2006;29(5):489–497. doi:10.1002/nur.20147.
32. Tavakol M, Dennick R. Making sense of Cronbach's alpha. *Int J Med Educ*. 2011;2:53–55. doi:10.5116/ijme.4dfb.8dfd.
33. Guest G, MacQueen KM, Namey EE. Introduction to applied thematic analysis. *Appl Them Anal*. 2012;3:20.
34. Ryan K, Gannon-Slater N, Culbertson MJ. Improving survey methods with cognitive interviews in small- and medium-scale evaluations. *Am J Eval*. 2012;33(3):414–430. doi:10.1177/1098214012441499.
35. Scala D, Menditto E, Armellino MF, et al. Italian translation and cultural adaptation of the communication assessment tool in an outpatient surgical clinic. *BMC Health Serv Res*. 2016;16(1):163. doi:10.1186/s12913-016-1411-9.
36. Norcini J, Anderson B, Bollela V, et al. Perrott V and others. Criteria for good assessment: consensus statement and recommendations from the Ottawa 2010 Conference. *Med Teach*. 2011;33(3):206–214. doi:10.3109/0142159X.2011.551559.
37. Swanson DB, van der Vleuten CP. Assessment of clinical skills with standardized patients: state of the art revisited. *Teach Learn Med*. 2013;25(supplement 1):S17–S25. doi:10.1080/10401334.2013.842916.

38. Epstein J, Santo RM, Guillemin F. A review of guidelines for cross-cultural adaptation of questionnaires could not bring out a consensus. *J Clin Epidemiol*. 2015;68(4):435–441. doi:10.1016/j.jclinepi.2014.11.021.
39. Ferranti DE, Makoul G, Forth VE, et al. Assessing patient perceptions of hospitalist communication skills using the Communication Assessment Tool (CAT). *J Hosp Med*. 2010;5(9):522–527. doi:10.1002/jhm.787.
40. Clayman ML, Bylund CL, Chewning B, et al. The impact of patient participation in health decisions within medical encounters: a systematic review. *Med Decis Making*. 2016;36(4):427–452. doi:10.1177/0272989X15613530.
41. Noble LM, Scott-Smith W, O'Neill B, et al. Consensus statement on an updated core communication curriculum for UK undergraduate medical education. *Patient Educ Couns*. 2018;101(9):1712–1719. doi:10.1016/j.pec.2018.04.013.
42. Morrow G, Rothwell C, Burford B, et al. Cultural dimensions in the transition of overseas medical graduates to the UK workplace. *Med Teach*. 2013;35(10):e1537–e1545. doi:10.3109/0142159X.2013.802298.
43. Kurtz S, Silverman J, Benson J, et al. Marrying content and process in clinical method teaching: enhancing the Calgary-Cambridge guides. *Acad Med*. 2003;78(8):802–809. doi:10.1097/00001888-200308000-00011.
44. Lisy K, Peters MD, Riitano D, et al. Provision of meaningful care at diagnosis, birth, and after stillbirth: a qualitative synthesis of parents' experiences. *Birth*. 2016;43(1):6–19. doi:10.1111/birt.12217.
45. Prinds C, Hvidt NC, Mogensen O, et al. Making existential meaning in transition to motherhood – a scoping review. *Midwifery*. 2014;30(6):733–741. doi:10.1016/j.midw.2013.06.021.
46. Silva de Sa MF, Romao GS. The need for a competence matrix in delivery rooms for the proper work of multiprofessional teams. *Rev Bras Ginecol Obstet*. 2019;41(1):1–3.
47. Cook DA, Brydges R, Ginsburg S, et al. A contemporary approach to validity arguments: a practical guide to Kane's framework. *Med Educ*. 2015;49(6):560–575. doi:10.1111/medu.12678.