

Teacher Behavior Checklist Content Validity According to Brazilian Teachers and Undergraduates

Marcelo H. O. Henklain¹

João dos Santos Carmo²

Verônica Bender Haydu³

Monalisa Muniz²

¹Universidade Federal de Roraima

²Universidade Federal de São Carlos

³Universidade Estadual de Londrina

Abstract

This research, divided into two studies, expanded the content validity examination of the Teacher Behavior Checklist (TBC). In Study 1, the objective was to identify the degree of relevance attributed by teachers and undergraduates to the 28 TBC items, based on what they consider to characterize an effective teacher. The participants of Study 1 were 85 teachers and 91 students from different Brazilian universities. They assessed the TBC items on a seven-point scale (ranging from “1 = totally irrelevant” to “7 = totally relevant”). Students (Mean = 5.7) and teachers (6.1) considered the TBC items to be relevant. Study 2 sought to identify the 10 most important TBC qualities based on student perceptions. The participants of Study 2 were 995 undergraduates whose choices of items corroborated results from previous TBC studies. These data encourage the use of TBC in teacher formative assessment and research on effective teaching.

Keywords: psychometrics; test validity; teacher performance assessment; higher education

Validade de Conteúdo do Teacher Behavior Checklist segundo Professores e Universitários Brasileiros

Resumo

Esta pesquisa, dividida em dois estudos, ampliou a investigação da validade de conteúdo do Teacher Behavior Checklist (TBC). No Estudo 1, o objetivo foi identificar o grau de relevância atribuído por professores e graduandos aos 28 itens do TBC, baseados no que consideram caracterizar um professor eficaz. Participaram 85 professores e 91 estudantes de diferentes universidades brasileiras. Eles avaliaram os itens do TBC segundo uma escala de setes pontos (“1 = totalmente irrelevante” a “7 = totalmente relevante”). Estudantes (Média = 5,7) e professores (6,1) consideraram os itens do TBC relevantes. O Estudo 2 buscou identificar quais são as 10 qualidades mais importantes do TBC segundo estudantes. Participaram 995 graduandos cujas escolhas de itens corroboraram resultados de estudos anteriores com o TBC. Esses dados incentivam o uso do TBC na avaliação formativa de professores e na pesquisa sobre ensino eficaz.

Palavras-chave: psicometria, validade do teste, avaliação de desempenho do professor, ensino superior

Validez de Contenido del Teacher Behavior Checklist Según Profesores y Universitarios Brasileños

Resumen

Esta encuesta, dividida en dos estudios, amplió la investigación de validez de contenido del Teacher Behavior Checklist (TBC). El estudio 1 tuvo por objetivo identificar el grado de relevancia atribuido por profesores y estudiantes a los 28 ítems del TBC, basados en lo que consideran caracterizar un profesor eficaz. Participaron 85 profesores y 91 estudiantes de diferentes universidades brasileñas. Ellos evaluaron los ítems del TBC según una escala de siete puntos (“1 = totalmente irrelevante” a “7 = totalmente relevante”). Estudiantes (Media = 5,7) y profesores (6,1) consideraron los ítems del TBC relevantes. El estudio 2 buscó identificar cuáles son las 10 cualidades más importantes del TBC según los estudiantes. Participaron 995 estudiantes cuyas elecciones de los ítems corroboraron resultados de estudios previos con el TBC. Estos datos incentivan el uso del TBC en la evaluación formativa de profesores y en la investigación sobre enseñanza eficaz.

Palabras clave: psicometría; validez del test; evaluación de desempeño del profesor; Enseñanza Universitaria

The Teacher Behavior Checklist (TBC) (Buskist, Sikorski, Buckley, & Saville, 2002) is an instrument recognized internationally by students, teachers, and researchers for its contributions to describing how an effective teacher¹ behaves in the classroom and, also, because it can be used to assist with formative evaluation of university teachers (Keeley, Smith, & Buskist, 2006). The purpose of this type of assessment is to help teachers to enhance their teaching practice and should not be used to guide administrative decisions, such as hiring, dismissal, promotion, etc. (Cassettari, 2014). Research into effective teaching and initiatives to develop resources for formative evaluation of higher education teachers in Brazil can be benefitted by using TBC.

In the study in which they developed the TBC, Buskist et al., (2002) asked 114 undergraduate students to list the qualities of those teachers in whose classes they learned a lot and in a pleasant way. This procedure gave rise to a list containing 47 qualities which was sent to a group of 184 undergraduates whose task was to identify three behaviors related to each of those qualities. The results were organized by three researchers into a list of 28 qualities and examples of corresponding behaviors.

That list was given to 916 undergraduates and 118 teachers who were asked to select the 10 main qualities of effective teachers. The objective was to identify whether there would be convergence between the participants as to what is fundamental for working as a teacher. Teachers and students agreed on six of the 10 qualities, differing only in the order of importance attributed to each one: teachers gave a higher rating to technical-pedagogical behaviors (e.g., effective communication), while students placed more value on teachers' accessible and confident interaction with their students. These results confirmed the

expectations of Buskist et al., (2002) about the convergence of opinions between teachers and students as to the most relevant items. Furthermore, the approach taken by their study was considered to be innovative because few instruments reported in the literature had the format of a checklist or had their items prepared based on the perspective of undergraduates as to what comprises an effective teacher. In addition, few instruments emphasized concrete teacher behaviors in the classroom.

A variety of studies followed, mainly to replicate the work of Buskist et al., (2002) using other samples. The first replication sought to investigate whether the results obtained by Buskist et al., were related to the type of university in which the instrument was administered. Schaeffer, Epting, Zinn, and Buskist (2003) asked 231 students and 99 teachers at a North American community college to select, based on TBC, the 10 main qualities of an effective teacher. They found results similar to Buskist et al., Agreement between teachers and students was found in 8 out of the 10 qualities: (a) approachable, (b) creative/interesting, (c) encouraging/caring, (d) enthusiastic, (e) flexible/open-minded, (f) knowledgeable, (g) realistic expectations/fair and (h) respectful. The differences in the answers once again showed that teachers place more value on technical/pedagogical qualities while relationship qualities were more important for students. These data indicated that it was possible, using TBC, to identify agreement between students and teachers regarding the main qualities of an effective teacher.

Despite its successful replication, remained the problem that although the samples studied were different from each other, they were predominantly composed of psychology undergraduates and teachers (Ford, 2016). This bias raised doubts regarding the extent to which the data obtained might be a product exclusive to those samples. Considering this bias, Ismail and Groccia (2017) compared the results of the 10 most important TBC items according to 448 teachers: 309 trained in the USA and 139 in other countries, mainly European and Asiatic countries. Convergence was found in 8 out of 10 qualities. Being knowledgeable about the subject matter and being enthusiastic are examples of qualities most listed in first or second place. Ford (2016), in turn, conducted a direct replication of the study undertaken by Buskist et al., (2002) with a sample of pharmacy students and teachers from Auburn University in the USA, thus broadening the diversity of participants in research on TBC. The

¹ The authors of TBC use expressions such as excellent teacher, master teacher, good teacher or effective teacher. In the literature these terms may or may not appear as synonyms. In our study we use the term effective teacher. An effective teacher creates conditions that favor the learning of behavioral objectives relevant for the learner and for society, without the stimuli associated with education becoming aversive for the learner during the teaching process (Henklain et al., in press). In order to be able to say that they are effective, these learned behaviors should be demonstrated by learners in natural settings, so as to alleviate or solve problem situations, and not only in the classroom (thus characterizing efficiency) (De Luca, 2013). The only exception to this rule were the instructions given to the participants, in which expressions such as excellent teacher or excellent teaching were used because they appeared to be easier to understand.

results showed agreement between students and teachers on six of the 10 qualities.

Among the set of researches based on what had been developed by Buskist et al., (2002), the study fulfilled by Jõemaa (2013) was the only one with the most divergent data about qualities most valued by students. This researcher showed the 28 TBC items to university students from different courses in Estonia and asked them to select the 10 main qualities of an effective teacher. The results showed that these students placed more value on a teacher's knowledge about the subject matter and that interpersonal interaction characteristics were not, on average, listed as being especially important, although younger participants did point to them as being among the 10 main qualities. This difference may be associated with the characteristics of this sample from Estonia and the diversity of the courses in which TBC was administered, a new aspect concerning preceding studies. As such, the degree of convergence between members of different samples about the perception of which TBC items are the most important continues to be an issue that requires further research. It is relevant to investigate this question because answering it may help to identify fundamental aspects (valid in teaching/learning settings in different cultures) regarding what comprises effective teaching.

A new step forward in the investigation of qualities of excellence was made by Buskist and Keeley (2018). This study assessed the extent to which data from several studies of TBC enable identification of "universal" qualities characterizing excellent teachers. They analyzed data from studies with teachers and students from the following countries: United States, Canada, Brazil, Colombia, Japan, China, Germany, Estonia, and Saudi Arabia. The results revealed two universally relevant quantities: "knowledgeable about the subject taught" (recognized in 100% of the studies with teachers and students) and "enthusiastic about teaching and the subject they teach" (recognized by 100% of the studies with teachers and 86% of studies with students). Although with less specific criteria, other qualities were also found to be universally relevant: "accessible/approachable" (teachers – 86%; students – 79%) and "effective communicator" (teachers – 86%; students – 79%). These data provide further encouragement for moving forward with research on TBC, since such research may indicate priorities for teacher training. Identifying priorities is especially important in countries like Brazil which is marked by cultural diversity and has demands in terms of teacher training and enhancement of education.

In addition to the set of studies described so far, other studies (Lammers, Savina, Skotko, & Churlyeva, 2010; Liu, Keeley, & Buskist, 2015) concerned themselves with investigating whether teachers and students from different cultures in teaching/learning settings considered each of the 28 TBC items to be relevant or characteristic of effective teachers. Lammers et al., for instance, assessed whether there was a difference between the pattern of answers given by university psychology teachers and students in the United States and those provided by teachers and students at two Russian universities. Their task was to assess each TBC item according to a scale of importance (ranging from "1 = not at all important" to "7 = extremely important") for the repertoire of an effective teacher. Study participants were 118 teachers and 179 students from the United States and 45 teachers and 222 students from Russia. No substantial differences were found in the participants' answers indicating that 21 of the TBC 28 items were important. Liu et al., in turn, asked Chinese psychology undergraduates to assess the frequency with which effective teachers showed TBC qualities, using a five-point scale, from "1 = never exhibit" to "5 = always exhibit". The Chinese students considered that all TBC qualities were frequent of effective teachers (mean greater than four), suggesting that despite cultural differences in teaching/learning settings, teaching practices comprising TBC are internationally recognized as being relevant.

The results found by Lammers et al., (2010) and Liu et al., (2015) are fundamental because they suggest that students and teachers, who are users and professionals directly affected by the use of TBC, did indeed identify these items as being relevant and characteristic of effective teaching. This result is evidence of TBC content validity which expands the significance of the results of the first set of studies regarding the 10 most important items and the convergence between participants. In that set of studies, participants could consider TBC to be an instrument of little relevance for effective teaching and, even so, could select the 10 most important items when the researcher requested them to do so. Taking the results of Lammers et al., and Liu et al., related to the first set of studies, there is evidence of convergence between teachers and students in relation to the 10 most important items in a set of 28 items where all of them are considered to be fundamental for comprising the repertoire of an effective teacher.

We found few studies in Brazil assessing convergence and what are the most important TBC items, or whether the items are relevant for excellent teaching. A search using exact terms conducted in June 2016 (and repeated in November 2017) of the *Biblioteca Digital Brasileira de Teses e Dissertações* and of the *Biblioteca Virtual de Psicologia do Brasil* using the descriptors “Teacher Behavior Checklist”, “*Checklist de Comportamentos do Professor*” and “*Lista de Comportamentos do Professor*”, found no results related to TBC.

By using other searching instruments, two adaptations of TBC in Brazil were found (Cacciari, 2015; Henklain et al., in press). The first adapted version was developed for university teachers to self-assess their performance and can be found in a study conducted by Cacciari. This adapted version of TBC was called *Lista de Comportamentos do Professor* (List of Teacher Behaviors) and has 19 items. Psychometric evidence was found for this version of the instrument.

The second adapted version arose from the Ph.D. dissertation of the lead author of this study and was prepared for students to assess the performance of their teachers (Henklain et al., in press). This study maintained the original 28-item TBC structure. The research was organized in the form of two studies to investigate diverse psychometric evidence. The objective of Study 1 was the adaptation of TBC for use by students to assess teachers. It should be emphasized that this study showed evidence of content validity (as assessed by three judges) as well as semantic validity to the instrument’s clearness (as assessed by 65 university students from a variety of courses in the Brazilian state of Roraima). The participants of Study 2 were 714 university students attending a public university in Roraima. Diverse evidence of validity was obtained, such as evidence of internal structure validity – verified by means of Exploratory Factor Analysis which led to a proposed two-factor model for the adapted version of TBC, which explains 49% of variance, this being a model similar to that found in the original version (see Keeley et al., 2006). Evidence of reliability was also observed through the test-retest method, the result of which was strong and significant ($r_s = 0.748$, $p < 0.01$, one-tailed, $N = 229$).

Further research was conducted by Henklain, Carmo, Haydu, and Muniz (2018), based on the Brazilian adaptation of TBC for students to assess teachers. This research was organized into two studies. In Study 1, they asked 186 psychology students and 76 psychology course teachers to assess each TBC item using a

seven-point scale, where “1 = totally irrelevant” and “7 = totally relevant”. The results showed that both teachers and students gave high mean scores for the TBC items (teachers: mean = 6.2, $SD = 0.55$; students: mean = 6, $SD = 0.63$), with teachers tending to rate relevance with higher scores than students ($U = 5608$, $p = 0.009$, $r = -0.16$, $N = 262$).

In Study 2, 467 psychology students were asked to indicate the 10 most important TBC items for an effective teacher. The data found showed 50% convergence with those of Buskist et al., (2002). The Brazilian students, as happened in previous studies with TBC, selected “knowledgeable about the subject matter” as the first and, therefore, most important quality.

Given existing knowledge and the potential of TBC to contribute to the formative evaluation of university teachers, the importance of expanding studies on this instrument in Brazil, samples in terms of undergraduate courses, since available data relates only to psychology courses. With these aspects in mind, we conducted the research here presented, which was divided into two studies. The objective of Study 1 was to identify the degree of relevance attributed by Brazilian university teachers and undergraduates to the 28 TBC items to what they considered as comprising the behavioral repertoire of an effective teacher. The objective of Study 2 was to identify which are the TBC 10 most important qualities among students, to compare this data with that of national and international studies. Type of comparison can assist with the identification of teaching practices considered necessary for an effective teacher in different cultural settings for teaching/learning.

Study 1

Method

Participants

Ninety-one university students from several different Brazilian public universities (79 students) and private universities (12 students) took part in Study 1. Fifty-two came from Brazil’s Northern region, three from the Northeast region, three from the Midwest region, 20 from the Southeast region and 13 from the Southern region. This sample was comprised of 52 females and 39 males, with a mean age of 23.6 years ($SD = 6.38$). The majority, 22%, had been taking the course for less than one semester (29.7% had completed seven semesters or more). Five students were taking

undergraduate courses in the area of Humanities, 16 in Health Sciences, 10 in Engineering, 33 in Applied Social Sciences, seven in Linguistics, Language/Literature and Arts, 15 in Exact and Earth Sciences, four in Biological Sciences and one in Agrarian Sciences.

Eighty-five university teachers also took part, 50 of whom were from public universities and 35 from private universities. Concerning the regions of Brazil where they taught at the time the study was conducted, 29 were from the North, six were from the Northeast, three from the Midwest, 29 from the Southeast and 18 from the South, 41 of whom were female, and 44 were male. Mean age was 39.9 years ($SD = 9.36$) and mean length of teaching experience was 10.9 years ($SD = 9.3$). In this sample, 10 teachers were from Humanities courses, 15 from Health Sciences, eight from Engineering, 24 from Applied Social Sciences, three from Linguistics, Language/Literature and Arts, 16 from Exact and Earth Sciences, four from Biological Sciences and five from Agrarian Sciences.

In this study the areas of knowledge were comprised of the following courses based on CAPES (Coordination for the Improvement of Higher Level Personnel/Ministry of Education) criteria: (a) Humanities – Psychology, Social Sciences, Education, Rural Education, Pedagogy, Special Education, History; (b) Health Sciences – Physical Education, Nursing, Physiotherapy, Gerontology, Medicine, Nutrition, Occupational Therapy; (c) Engineering – Environmental Engineering, Electrical Engineering, Civil Engineering, Mechanical Engineering, Materials Engineering, Mechatronics Engineering, Forestry Engineering, Production Engineering, Chemical Engineering and Control and Automation Engineering; (d) Applied Social Sciences – Administration, Architecture and Urbanism, Industrial Design, Law, Economics, Business Management, Hospital Management, Indigenous Territory Management, Journalism, Publicity and Advertising and Social Services; (e) Linguistics, Languages/Literature and Arts – Visual Arts, Languages/Literature and Music; (f) Exact and Earth Sciences – Computer Science, Information Systems, Chemistry, Physics, Statistics, Systems Analysis and Development Technologist, Mathematics and Geology; (g) Biological Sciences – Biology; (h) Agrarian Sciences – Agronomy, Veterinary Medicine and Zootechnics.

Instruments

The online protocol was comprised of: (a) TBC adapted for use by Brazilian university students to

assess teachers (28 item and seven-point relevance scale, ranging from “1 = totally irrelevant” to “7 = totally relevant”); and (b) Student Characterization Questionnaire. This questionnaire requested information on whether the respondent was a university teacher or an undergraduate, their higher education institution, city, and state of the institution, name of the undergraduate course (studied or taught depending on the type of participant), gender, age and level of education. Teachers were also asked how long they had been teaching, and students were asked how long they had been taking their undergraduate course.

Procedure

The sample was defined by convenience. The study was announced in the following social media: Facebook, LinkedIn, Twitter and WhatsApp (highlighting that only university teachers or university students could take part). Electronic messages were also sent to the contacts of the lead author. Posters presenting the study were affixed, and pamphlets were distributed to university students at two universities, one in Roraima and the other in Paraná. As a final strategy, a Facebook page was created to push a publication about the study to Brazilian university teachers and students.

When participants accessed the online study form, the first thing they saw was a Free and Informed Consent form approved by the Federal University of Roraima (UFRR) Human Research Ethics Committee, Certification of Submission for Ethical Appraisal (CAAE) No. 54448416.6.0000.5302. They had to read the Free and Informed Consent form and, if they agreed to participate, they had to confirm that they agreed to the conditions of the study. The announcements and posters, the Free and Informed Consent form and the first item of the instrument itself all explained that participants had to be Brazilian, university teacher or university student. The participants' task was to assess the degree of relevance of each TBC item regarding what they considered to be important for the profile of an excellent teacher. This concept was explained to participants as being a profile that favors a pleasant learning experience. Data collection took place between April and June 2017.

Data analysis

The software we used to analyze the data was R Studio based on R version 3.4.1 (R Development Core Team, 2017) together with the Psych package (Revelle, 2017), the Car package (Fox & Weisberg, 2011) and the Coin package (Hothorn, Hornik, van de Wiel, &

Zeileis, 2006). Using the Shapiro-Wilk normality test (5% significance level), we found that data distribution was not normal (all items had significant results, p -value < 0.05). For this reason, we used nonparametric statistics in the data analyses.

We calculated the mean scores for each participant's data and those of the groups of teachers and students, as well as the respective standard deviations of the groups. Total Type 1, 2 and 3 scores (corresponding to levels of irrelevance) were quantified and classified as *Irrelevant*. Total Type 4 scores (neither irrelevant nor relevant) were calculated and classified as *Moderate*. Total Type 5, 6 and 7 scores (corresponding to levels of relevance) were calculated and classified as *Relevant*. These data were used to show response patterns in the most straightforward manner and were also used to calculate the Content Validity Ratio (CVR), as proposed by Pacico and Hutz (2015) as a measurement of content validity. CVR results greater than or equal to 0.29 (in the case of this study in which there were more than 40 participants) were considered to be acceptable, while results lower than this indicated insufficient agreement between the participants concerning to the relevance of each item. We also calculated the mean score and the standard deviation of each item for the student and teacher samples.

Finally, we recognized that even though there might be similarity and agreement, there could be a difference in the degree of irrelevance or relevance attributed to the TBC items by teachers and students. To assess this, we added the scores given by each participant to the Irrelevant, Moderate and Relevant categories which produced one score per participant. Using this data we performed the Mann-Whitney two-tailed nonparametric test to assess differences between two independent samples, taking $p < 0.05$ as the criterion for accepting the alternative hypothesis. We performed Levene Test and calculated the Variance Ratio (the result had to be less than two, as suggested by Field, 2009) to assess variance homogeneity between the groups of teachers and students. We also calculated Cronbach's alpha.

Results

The scale's alpha was equal to 0.94 (95% confidence interval = 0.92 – 0.95), which is considered excellent (George & Mallery, 2002) and evidence of the reliability of the TBC version used in this study. The 91 university students selected Levels 1 to 3 of the scale (irrelevant)

186 times, Level 4 (moderate) 250 times and Levels 5 to 7 (relevant) 2112 times. The 85 university teachers selected Levels 1 to 3 (irrelevant) only 89 times, Level 4 (moderate) 121 times, and Levels 5 to 7 (relevant) 2170 times. In the sample as a whole, irrelevant levels were selected 275 times, the moderate level 371 times and relevant levels 4282 times. These data demonstrate that the participants tended to consider relevant the TBC items as components of the effective teacher's repertoire. Table 1 shows the overall results and the results by type of participant for the CVR, the mean scores of each item and the standard deviations.

Table 1 shows that for students Items 3, 12 and 17 had the lowest mean scores (4.9, 4.7 and 4.9, respectively) and CVR results (0.21, 0.19 and 0.3) below or very close to the criterion of 0.29. Despite their being the lowest, these three mean scores fall between the moderate and relevant levels. As for the teachers, no item had a mean score below five or an inadequate CVR value. When analyzing the overall data, there is also neither a mean score below five or an inadequate CVR value. In general, when analyzing the CVR means and the mean scores per item, it can be seen that the results are high and adequate both for teachers and for students. This data, therefore, is an evidence of content validity obtained from those who could be TBC users in applied settings.

Although students and teachers tended to assess all TBC items as relevant, we felt it was important to analyze whether there was a difference in the degree of relevance identified by these participants. As can be seen in Table 1, students had a mean score of 5.7 ($SD = 0.9$; Median = 5.96) on the relevance scale, while for teachers it was 6.1 ($SD = 0.6$; Median = 6.18). When performing the Mann-Whitney test, we found a statistically significant difference between teachers and students, whereby teachers assessed the TBC items more positively ($U = 2744$, $p = 0.00086$, $r = -0.251$, $N = 176$). Levene's test indicated the existence of variance homogeneity between the groups, $F(1, 174) = 2.008$, $p = 0.16$, although the variance ratio was equal to 2.26, thus exceeding the criterion suggested by Field (2009).

Study 2

Method

Participants

A total of 995 students were enrolled, 525 female, 470 male, mean age of 24 years ($SD = 6.3$). In this

Table 1.
CVR, Mean and Standard Deviation – Overall and By Type of Participant.

Item	Student			Teacher			Overall		
	CVR	M	SD	CVR	M	SD	CVR	M	SD
1. Accessible	0.74	5.8	1.3	0.88	5.9	1.0	0.81	5.9	1.1
2. Approachable	0.60	5.7	1.4	0.98	6.5	0.8	0.78	6.1	1.2
3. Authoritative	0.21	4.9	1.3	0.58	5.7	1.5	0.39	5.3	1.5
4. Confident	0.89	6.2	1.0	0.98	6.6	0.7	0.93	6.4	0.8
5. Creative and Interesting	0.74	5.9	1.3	0.91	6.3	0.9	0.82	6.1	1.1
6. Effective Communicator	0.82	6.1	1.3	0.91	6.6	1.0	0.86	6.3	1.2
7. Cares for Students	0.54	5.5	1.3	0.88	6.2	0.9	0.70	5.8	1.2
8. Enthusiastic	0.74	6.0	1.3	0.93	6.6	0.7	0.83	6.2	1.1
9. Establishes Goals	0.67	5.8	1.3	0.91	6.3	0.9	0.78	6.0	1.1
10. Flexible	0.69	5.6	1.4	0.91	6.1	1.1	0.80	5.8	1.2
11. Good Listener	0.69	5.7	1.2	0.93	6.4	0.9	0.81	6.0	1.1
12. Positive Attitude	0.19	4.7	1.5	0.44	5.3	1.8	0.31	5.0	1.7
13. Humble	0.67	5.7	1.5	0.79	6.1	1.3	0.73	5.9	1.4
14. Knowledgeable	0.89	6.4	1.1	0.98	6.6	0.7	0.93	6.5	0.9
15. Prepared	0.89	6.3	1.1	0.88	6.4	0.9	0.89	6.3	1.0
16. Presents Current Information	0.80	6.1	1.3	0.88	6.2	1.0	0.84	6.2	1.2
17. Professional	0.30	4.9	1.8	0.39	5.3	1.6	0.34	5.1	1.7
18. Promotes Class Discussion	0.54	5.3	1.6	0.84	6.0	1.0	0.68	5.7	1.4
19. Promotes Critical Thinking	0.71	5.8	1.3	0.88	6.3	1.0	0.80	6.1	1.2
20. Provides Constructive Feedback	0.80	6.1	1.2	0.84	6.3	1.1	0.82	6.2	1.2
21. Punctuality	0.82	6.0	1.2	0.84	6.0	1.1	0.83	6.0	1.2
22. Rapport	0.34	5.1	1.5	0.53	5.5	1.5	0.43	5.3	1.5
23. Realistic Expectations	0.78	5.9	1.3	0.84	6.1	1.2	0.81	6.0	1.2
24. Respectful	0.87	6.4	1.1	0.98	6.7	0.6	0.92	6.5	0.9
25. Sensitive and Persistent	0.65	5.6	1.5	0.91	6.1	0.8	0.77	5.9	1.3
26. Strives to Be a Better Teacher	0.74	6.0	1.3	0.88	6.4	1.0	0.81	6.2	1.2
27. Technologically Competent	0.49	5.4	1.4	0.72	5.8	1.4	0.60	5.6	1.4
28. Understanding	0.60	5.6	1.4	0.69	5.8	1.3	0.65	5.7	1.3
Mean	0.7	5.7	1.3	0.8	6.1	1.1	0.7	5.9	1.2

Note. CVR = Content Validity Ratio; M = Mean; SD = Standard Deviation.

sample, 793 were from a public university in the state of Roraima (Northern region of Brazil), 27 were from two private universities also in Roraima, 15 were from a public university in the state of Paraná (Southern region of Brazil) and 160 were from a public university in the state of São Paulo (Southeast region of Brazil). Distribution according to the area of knowledge was

as follows: Applied Social Sciences, 257; Engineering, 190; Humanities, 147; Agrarian Sciences, 127; Exact and Earth Sciences, 86; Biological Sciences, 61; Linguistics, Languages/Literature and Arts, 54; Health Sciences, 44; Other, 29.

In this study the areas of knowledge were comprised of the following courses based on CAPES

criteria: (a) Humanities – Philosophy, Geography, History, Special Education, Pedagogy and International Relations; (b) Health Sciences – Physical Education, Nursing, Physiotherapy and Dentistry; (c) Engineering – Civil Engineering, Electric Engineering, Mechanical Engineering and Chemical Engineering; (d) Applied Social Sciences – Administration, Architecture and Urbanism, Accountancy, Law, Economics and Journalism; (e) Linguistics, Languages/Literature and Arts – Visual Arts, Languages/Literature and Music; (f) Exact and Earth Sciences – Computer Science, Physics, Geology, Computing, Mathematics, Chemistry and Information Systems; (g) Biological Sciences – Biology, Biotechnology; (h) Agrarian Sciences – Agronomy and Zootechnics; (i) Other – Executive Secretary.

Instruments

The protocol used was comprised of: (a) an item instructing participants to write the number and the name of the 10 most important TBC qualities (participants had access to the TBC items adapted for Brazilian students); (b) questions to characterize the participants.

Procedure

We made contact with teachers from the different institutions by e-mail or WhatsApp and requested 40 minutes of lesson time to administer a questionnaire about teacher qualities. As agreed beforehand, when arriving at the classroom the researcher asked the teacher to leave the classroom while data were collected and explained the objectives of the research to the students. The participant's task consisted of selecting the ten TBC qualities they considered to be most important for an effective teacher, i.e., a teacher who favors a pleasant learning experience for students.

This study was approved by the Federal University of Roraima (UFRR) Human Research Ethics Committee, Certification of Submission for Ethical Appraisal (CAAE) No. 54448416.6.0000.5302. All those who took part in the study signed the Free and Informed Consent form.

Data analysis

We calculated the number of times each quality was selected by the participants and we created a ranking comprised as follows: the most selected quality came in 1st position, the second most selected quality came in 2nd position, and so on. If two or more

qualities received the same rating, they were given the same number. The one assigned to the next most selected quality was equivalent to the number assigned to the tied qualities plus the times that this number was repeated. Therefore, if two qualities came in 5th position, the next quality would be the 5th position + 2, i.e., 7th position.

Results

Table 2 shows the number of times students selected each TBC quality and its corresponding ranking exhibits that the 10 most selected items in descending order were: “14. Knowledgeable about the subject matter”, “1. Accessible”, “26. Strives to be a better teacher”, “15. Prepared”, “6. Effective communicator”, “5. Creative and interesting”, “8. Enthusiastic”, “24. Respectful”, “28. Understanding” and “2. Approachable”. The least selected item was “25. Sensitive and persistent”, having only 94 indications to comprise the list of the 10 main items.

Discussion

The objectives of this research were to identify the degree of relevance attributed by teachers and students from different courses to the TBC items (Study 1), and to identify the 10 most important qualities among college students (Study 2). In Study 1 we found that the overall mean value for CVR was 0.7. This value is high and suggests the existence of agreement regarding the relevance of the TBC items. This data on the convergence between people who face different contingencies of the teaching and learning process (e.g., demands, functions, working conditions, etc.), suggests that the TBC items do represent relevant qualities for characterizing the practices of an effective teacher (Buskist & Keeley, 2018). The data, therefore, indicate what students expect of their teachers and what teachers expect of themselves. As such, when teachers demonstrate these qualities, they tend to be assessed positively by their students. This relationship favors, for example, the forming of ties with them, increasing the likelihood of students paying attention to what these teachers say, doing what they ask, and so on (Benson, Cohen, & Buskist, 2005).

None of the aspects above is a guarantee of learning, but they can favor it. It tends to be less likely, although not impossible, that students learn with teachers they do not like, do not pay attention to what

Table 2.
Number of Items Selected and Ranking.

Item	Total selections	Ranking
14. Knowledgeable	681	1
1. Accessible	612	2
26. Strives to Be a Better Teacher	508	3
15. Prepared	470	4
6. Effective Communicator	464	5
5. Creative and Interesting	449	6
8. Enthusiastic	446	7
24. Respectful	425	8
28. Understanding	416	9
2. Approachable	410	10
7. Cares for Students	404	11
10. Flexible	395	12
19. Promotes Critical Thinking	394	13
4. Confident	358	14
13. Humble	326	15
21. Punctuality	326	15
9. Establishes Goals	322	17
12. Positive Attitude	291	18
18. Promotes Class Discussion	282	19
11. Good Listener	281	20
16. Presents Current Information	271	21
23. Realistic Expectations	241	22
27. Technologically Competent	233	23
3. Authoritative	231	24
20. Provides Constructive Feedback	223	25
22. Rapport	223	25
17. Professional	174	27
25. Sensitive and Persistent	94	28

they say, do not do what they ask or do not tolerate shortcomings they may have. Moreover, if teachers themselves consider these TBC items to be important, it can be assumed that being successful in demonstrating the qualities these items refer to is important to generate self-confident feelings about the work they do and that they have more reasons to persist with their work and further improve themselves.

The data on the overall mean score for the relevance of the TBC items are evidence for the interpretations mentioned above. We found a mean of 5.9

($SD = 0.4$) scores for relevance. This value is high, given that the maximum score on the scale was seven. The data do support the assertion that for this Brazilian sample TBC is a relevant source of information about good teaching practices.

Despite the convergence between the participants' data, we found a difference in the degree of relevance attributed to the TBC items, namely: teachers showed a tendency to attribute higher scores of relevance to the TBC items when compared to students. This data is especially relevant for defending the

use of TBC to assist formative evaluation processes (Keeley et al., 2006), given that teachers who will be evaluated recognized the relevance of the instrument.

The data relating to the mean values for the relevance of the TBC items corroborate findings of previous studies, such as those conducted by Lambers et al., (2010), Liu et al., (2015) and Henklain et al., (2018). The study conducted by Henklain et al., for example, obtained mean relevance scores of 6 and 6.2 for teachers and students, respectively. That study by Henklain et al., also found a statistically significant difference in that teachers gave higher scores than students for the TBC items.

In Study 2, the results we obtained (the selection of Items 14, 1, 26, 15, 6, 5, 8, 24, 28 and 2, as shown in Table 2) corroborate the findings described in the literature. For example, in the study conducted by Henklain et al., (2018) with psychology undergraduates, the 10 most selected items were: 14, 1, 19, 8, 26, 10, 6, 7, 5 and 15. In this case, there were seven agreements between our study and that of Henklain et al., Buskist et al., (2002), in turn, identified in their sample of United States psychology students the following 10 most relevant TBC items: 23, 14, 28, 2, 24, 5, 12, 7, 10 and 8. Concerning Buskist et al., we found six corresponding items in our study. About the data obtained by Schaeffer et al., (2003) with students from different courses and different types of colleges in the United States (14, 2, 23, 24, 5, 12, 8, 7, 10 and 28), we also found six corresponding items in our study. When compared to the data collected by Ford (2016) from United States pharmacy undergraduates, 14, 6, 23, 2, 8, 24, 4, 7, 28 and 1, we found seven corresponding items in our study. Concerning the data obtained by Jõemaa (2013) with students from different courses in Estonia (14, 8, 20, 2, 5, 17, 23, 16, 15 and 10), we found five corresponding items in our study. The data suggest the existence of a pattern in the selection of certain qualities as being the main ones comprising the repertoire of an effective teacher, although the data from the study conducted by Jõemaa are not as close to the pattern found in data collected in Brazil and the USA. Finally, about the data obtained by Buskist and Keeley (2018), we found that the items they identified as being “universally” important, “knowledgeable about the subject matter” (Item 14), “enthusiastic about teaching and the subject taught” (Item 8), “approachable” (Item 2) and “effective communicator” (Item 6), were all included in the results of our sample.

Despite the convergence found regarding the relevance of the TBC items and the pattern in the selection of the most important items, it must be emphasized that data may vary depending on the sample (even in the same country) and the type of task the participants are asked to perform. For example, in Study 1, the items with the highest mean scores were: 14, 24, 4, 6, 15, 8, 16, 20, 26, and 2 (see Table 1). There was convergence between seven of these items, and the most selected in Study 2. Although this is a high convergence we didn't find a perfect agreement. In other words, there is a tendency for differences to occur between samples.

It is noteworthy that in general Item 14, “Knowledgeable” is highly recognized or selected in studies with TBC. This pattern suggests that adequate qualification about what the teacher teaches should be a priority aspect in his/her training and evaluation. The valuing of the Item 14 can be used to rethink the criteria for attributing disciplines to teachers. At times this decision can be made based more on need or convenience and less based on the teacher's effective preparation.

Although results are positive, new researches are needed to deal with the limitations of our study. Such studies should, for example, expand samples in terms of quantity and diversity (e.g., greater balance between participation of public and private universities, students from different courses and from different semesters, teachers with greater and lesser time of teaching service, etc.). It is also relevant to define the participants in a probabilistic manner in order to increase the possibility of generalizing the results. Specifically, concerning Study 2, we suggest that future research attempt to control a possible order effect arising from the fact that the TBC items are always presented in an order going from 1 to 28. Would the results be different if they were presented in decreasing order, for instance? Or in other words, is there a tendency for participants to neglect the items at the end of the instrument, for example?

The data presented support for the continuity of research on TBC and its use in formal educational settings. Moreover, our data replicate results obtained in other countries to the agreement between teachers and students regarding the relevance of the TBC items and regarding which items are most important. The identification of shared priorities in terms of teacher qualities in samples from different cultural settings for teaching/learning means it is possible

to pinpoint behaviors which should be prioritized to favor excellence in teaching. Furthermore, TBC helps to indicate grounds for teacher training and evaluation that may be useful in different countries. It also shows that teaching knowledge and experiences can be shared since, certain teaching practices are perceived as relevant by students from different cultures. All of this is important for organizations such as The United Nations Educational, Scientific and Cultural Organization (UNESCO; Delors, 2010) concerned with Education worldwide, as well as for educators and psychologists seeking to understand what is fundamental for promoting learning.

References

- Benson, T. A., Cohen, A. L., & Buskist, W. (2005). Rapport: Its relation to students attitudes and behaviors toward teachers and classes. *Teaching of Psychology, 32*(4), 237-239. doi: 10.1207/s15328023top3204_8
- Buskist, W., & Keeley, J. W. (2018). Searching for universal principles of excellence in college and university teaching. *New Directions for Teaching and Learning, 156*, 95-105. doi: 10.1002/tl.20321
- Buskist, W., Sikorski, J., Buckley, T., & Saville, B. K. (2002). Elements of master teaching. In S. F. Davis & W. Buskist (Orgs.), *The teaching of psychology: Essays in honor of Wilbert J. McKeachie and Charles L. Brewer* (pp. 30-39). New York: Psychology Press.
- Cacciari, M. B. (2015). Employment or vocation? Correlates of well-being in academics [Emprego ou vocação? Correlatos do bem-estar em professores universitários] (Master's Degree Thesis, Psychology Masters Program, Federal University of Espírito Santo). Retrieved from <https://goo.gl/JK8Ky9>
- Cassettari, N. (2014). Teacher assessment: A matter of choice [Avaliação de professores: Uma questão de escolhas]. *Estudos em Avaliação Educacional, 25*(57), 166-197. doi: 10.18222/eaec255720142829
- De Luca, G. G. de (2013). Effectiveness assessment of a contingency based program to develop behaviors constituting the general class "assessing information reliability" [Avaliação da eficácia de um programa de contingências para desenvolver comportamentos constituintes da classe geral "avaliar a confiabilidade de informações"] (Ph.D. Dissertation, Postgraduate Psychology Program, Federal University of Santa Catarina). Retrieved from <https://bit.ly/2GYXL9C>
- Delors, J. (Org.). (2010). Learning: the treasure within; report to UNESCO of the International Commission on Education for the Twenty-first Century. Brasília: UNESCO. Retrieved from <http://bit.ly/2vMB8SO>
- Field, A. (2009). *Discovering Statistics Using IBM SPSS Statistics* [Descobrimos a estatística usando o SPSS]. Porto Alegre: Artmed.
- Ford, C. R. (2016). *Identifying effective teaching behaviors of pharmacy faculty master teachers* (Ph.D. Dissertation, Auburn University Graduate School). Retrieved from <https://goo.gl/Y2PbEa>
- Fox, J., & Weisberg, S. (2011). An {R} companion to applied regression. Thousand Oaks CA: Sage. Retrieved from <https://goo.gl/qaaADJ>
- George, D., & Mallery, P. (2002). *SPSS for Windows step by step: A simple guide and reference*. Boston: Allyn & Bacon.
- Henklain, M. H. O., Carmo, J. S., Haydu, V. B., & Muniz, M. (2018). Brazilian faculty and student perspectives on excellent teaching. *New Directions for Teaching and Learning, 156*, 31-39. doi: 10.1002/tl
- Henklain, M. H. O., Carmo, J. S., Haydu, V. B., Muniz, M., Buskist, W., & Keeley, J. W. (in press). Teacher Behavior Checklist: Psychometric evidence in teacher evaluation by Brazilian college students. *Paideia*.
- Hothorn, T., Hornik, K., van de Wiel, M. A., & Zeileis, A. (2006). A Lego system for conditional inference. *The American Statistician, 60*(3), 257-263. doi: 10.1198/000313006X118430
- Ismail, E. A., & Groccia, J. E. (2017). Foreign and U.S.-educated faculty members' views on what constitutes excellent teaching: Effects of gender and discipline. *To Improve the Academy, 36*(1), 20-38. doi:10.1002/tia2.20056
- Jõemaa, K. (2013). *Student perceptions of master teacher in Estonian universities* (Master's Degree Thesis, University of Tartu, Estonia). Retrieved from <https://goo.gl/YYcSz5>
- Keeley, J., Smith, D., & Buskist, W. (2006). The Teacher Behaviors Checklist: Factor analysis of its utility

- for evaluating teaching. *Teaching of Psychology*, 33(2), 84-91. doi: 10.1207/s15328023top3302_1
- Lammers, W. J., Savina, E., Skotko, D., & Churlyayeva, M. (2010). Faculty and student perceptions of outstanding university teachers in the USA and Russia. *Educational Psychology: An International Journal of Experimental Educational Psychology*, 30(7), 803-815. doi: 10.1080/01443410.2010.512382
- Liu, S., Keeley, J., & Buskist, W. (2015). Chinese college students' perceptions of characteristics of excellent teachers. *Teaching of Psychology*, 42(1), 83-86. doi: 10.1177/0098628315620888
- Pacico, J. C., & Hutz, C. S. (2015). Validity [Validade]. In C. S. Hutz, D. R. Bandeira, & C. M. Trentini (Orgs.), *Psicometria* (pp. 71-84). Porto Alegre: Artmed.
- R Development Core Team (2017). R: A language and environment for statistical computing. *R Foundation for Statistical Computing*. Retrieved from <https://www.R-project.org/>
- Revelle, W. (2017). Psych: Procedures for personality and psychological research. *Northwestern University*. Retrieved from <https://goo.gl/oxG99x>
- Schaeffer, G., Epting, K., Zinn, T., & Buskist, W. (2003). Student and faculty perceptions of effective teaching: A successful replication. *Teaching of Psychology*, 30(2), 133-136. Retrieved from <https://goo.gl/7BGLmE>

Recebido em: 29/05/2018

Reformulado em: 24/02/2019

Aceito: 24/03/2019

About the authors:

Marcelo H. O. Henklain is a psychologist graduated at Universidade Estadual de Londrina, PhD in Psychology at Universidade Federal de São Carlos, Associate Professor of Psychology (Universidade Federal de Roraima), and external collaborator of the National Institute of Sciences and Technology on Behavior, Cognition and Teaching (INCT-ECCE).

ORCID: <https://orcid.org/0000-0001-9884-8592>

E-mail: marcelo.henklain@ufr.com.br

João dos Santos Carmo is a psychologist graduated at Universidade Federal do Pará, PhD in Education at Universidade Federal de São Carlos, Associate Professor at the Department of Psychology of the Universidade Federal de São Carlos, and researcher at the National Institute of Sciences and Technology on Behavior, Cognition and Teaching (INCT-ECCE).

ORCID: <https://orcid.org/0000-0003-3913-7023>

E-mail: jcarmo@ufscar.br

Verônica Bender Haydu is a psychologist graduated at Universidade Estadual de Londrina, PhD in Psychology at Universidade de São Paulo, Associate Professor at Universidade Estadual de Londrina. Scientific consultant of the Tesla Institute of Sciences and Technology.

ORCID: <https://orcid.org/0000-0002-4522-8338>

E-mail: jcarmo@ufscar.br

Monalisa Muniz is a psychologist, with a Masters and PhD in Psychological Evaluation at Universidade São Francisco, Professor at the Department of Psychology of the Universidade Federal de São Carlos (at the undergraduate and stricto sensu graduate levels) and researcher at the Human Development and Cognition Laboratory.

ORCID: <https://orcid.org/0000-0003-1628-6296>

E-mail: monamuniz@gmail.com

Contact:

Av. Cap. Ene Garcez, 2413, Aeroporto
Boa Vista-RR, Brasil
CEP: 69310-000