

Revista da
ASSOCIAÇÃO MÉDICA BRASILEIRA

www.ramb.org.br



Original article

Proposal for a unified selection to medical residency programs[☆]

Sônia Ferreira Lopes Toffoli^{a,b,*}, Olavo Franco Ferreira Filho^c, Dalton Francisco de Andrade^d

^aGraduate Program in Production Engineering, Universidade Federal de Santa Catarina, Florianópolis, SC, Brazil

^bUniversidade Estadual de Londrina, Londrina, PR, Brazil

^cDepartment of Internal Medicine, Universidade Estadual de Londrina, Londrina, PR, Brazil

^dDepartment of Informatics and Statistics, Universidade Federal de Santa Catarina, Florianópolis, SC, Brazil

ARTICLE INFO

Article history:

Received 15 February 2013

Accepted 30 June 2013

Keywords:

Medical residency

Selection process

Classical Test Theory

Item Response Theory

A B S T R A C T

Objective: This paper proposes the unification of entrance exams to medical residency programs (MRP) in Brazil. Problems related to MRP and its interface with public health problems in Brazil are highlighted and how this proposal is able to help solving these problems.

Methods: The proposal is to create a database to be applied in MRP unified exams. Some advantages of using the Item Response Theory (IRT) in this database are highlighted.

Results: The MRP entrance exams are developed and applied decentralized where each school is responsible for its examination. These exams quality are questionable. Reviews about items quality, validity and reliability of appliances are not common disclosed.

Conclusion: Evaluation is important in every education system bringing on required changes and control of teaching and learning. The proposal of MRP entrance exams unification, besides offering high quality exams to institutions participants, could be as an extra source to rate medical school and cause improvements, provide studies with a database and allow a regional mobility.

© 2013 Elsevier Editora Ltda. All rights reserved.

Proposta de seleção unificada aos Programas de Residência Médica

R E S U M O

Palavras-chave:

Residência médica

Processo seletivo

Teoria Clássica do Teste

Teoria de Resposta ao Item

Objetivo: Este artigo propõe a unificação dos exames de acesso aos programas de Residência Médica (RM) no Brasil. São destacados problemas relacionados à RM e a sua interface com problemas da saúde pública no Brasil e como esta proposta pode auxiliar no enfrentamento desses problemas.

Métodos: A proposta consiste na criação de um banco de itens para ser aplicado na seleção unificada para a RM. São destacadas algumas vantagens em utilizar a Teoria de Resposta ao Item (TRI) nesse banco de itens.

[☆]Study conducted at the Universidade Estadual de Londrina, Londrina, PR, Brazil.

*Corresponding author.

E-mail: sonialopes@uel.br (S.F.L. Toffoli).

Resultados: Os exames de seleção para os programas de RM são elaborados e aplicados de forma descentralizadas, cada instituição é responsável por sua avaliação. A qualidade dessas provas é questionável, os estudos referentes à qualidade dos itens, a validade e a confiabilidade dos instrumentos não são comumente divulgados.

Conclusão: A avaliação é importante em todo sistema de ensino, provocando transformações necessárias e monitoramento do ensino e da aprendizagem. A proposta da unificação da prova de seleção para a RM, além de oferecer uma avaliação de qualidade elevada às instituições participantes, poderia servir como mais um recurso para avaliar e consequentemente provocar intervenções de melhorias nos cursos de graduação em medicina, fornecer dados para estudos e permitir uma mobilidade regional.

© 2012 Elsevier Editora Ltda. Todos os direitos reservados.

Introduction

Medical Residency (MR) is a program of in-service training considered the best way to train physicians for professional practice. Currently, in some specialties, there is a discrepancy between the number of applicants and the number of MR vacancies, which results in fierce competition in the selection processes, especially in the more developed regions of Brazil.^{1,2} Although institutionalized for over 30 years in Brazil, MR lacks formulation of specific policies and mechanisms for coordination between the need and supply of professional training.^{3,4} Complying with this demand, the Programa Nacional de apoio à formação de médicos especialistas em áreas estratégicas (Pro-Residência) was launched in 2009, in order to favor the training of specializing physicians according to the needs of the Brazilian Unified Health System⁵ (Sistema Único de Saúde - SUS).

Studies approach the distribution of MR vacancies in different regions of Brazil, showing a concentration of programs and institutions that offer MR in the Southeast and South regions far superior than what is offered in the North and Northeast regions, but also that the percentage of scholarships offered in these disadvantaged regions has increased after the Pró-Residência program.¹

Policies to encourage retention of physicians in underserved places are still being used worldwide. These incentives can be financial or not – although important, when not accompanied by other benefits, the salary increment loses its strength.^{6,7} Examples of non-financial incentives are good housing conditions and schooling for children, continuing professional education options, possibility of career advancement, inclusion of assistance programs to residents as a preventive measure to stress and anxiety, and differentiated competition conditions for vacancies in MR.⁷⁻¹⁰

Among other proposals suggested by experts as complementary actions are the implementation of new MR programs in less privileged regions, refinancing educational debts and granting scholarships to students and residents in exchange for medical practice in underserved areas, the creation of a national career and proposing changes in the entrance

examinations of MR.^{6,7,10,11} The policies adopted to solve or alleviate the shortage of health professionals in barren areas are varied, but among them actions related to medical education are always mentioned, particularly with MR.

Currently, experts are concerned with improving the health care system to provide more efficient and humanistic services. There are several studies, debates and meetings on professional training and its integration into the health system. The National Meeting of Medical Entities (*Encontro Nacional de Entidades Médicas*) in 2010¹² resulted in some proposals concerning the formation of Medical Residents, among which: the guarantee of a MR vacancy for each student who concludes the undergraduation, expansion of MR vacancies according to demand from the public health sector, unifying the selection criteria of MR tests, minimizing the subjectivity in the assessment of inflow of residents and non-prioritization of the local egress over candidates from outside.

Fully meeting the majority of these proposals, and adding it to the current policies, this research proposes an unified selection exam throughout the country in order to change the access model to MR programs, adding consistent advantages with current actions aimed at training physicians and guiding these professionals to priority areas in SUS.

The uneven geographical distribution of physicians is considered serious since it generates shortages of professionals in remote areas. In addition to the shortage of physicians to work in primary health care, there are few medical specialists.^{1,5,7,13} In the article *Distribuição de Vagas de Residência Médica e de Médicos nas Regiões do País*, after citing examples from various studies around the world, Nunes et al.,² categorically state that the supply of quality residence helps to fixate physicians in the regions in which they attended such programs, and that the MR – not the undergraduation – is largely responsible for redistributing physicians.

The proposal for unified selection processes with a similar model to the National Secondary Education Examination¹⁴ (ENEM, acronym in Portuguese) could contribute to tackling the problem of lack of assistance in underprivileged areas. In such models, the candidates ranked on a common scale could be directed, respecting their preferences, to far locations, meeting

the current public policies to guide specialists to underserved regions.

The unified process for MR vacancies in Brazil would be done through a bank of items developed according to criteria of validity and reliability, described by current measure theories, in which the candidates would be ranked on a common scale.

As vacancies in the most competitive institutions are filled, the candidates are distributed to other institutions, according to their rank. It is believed that due to some advantages, many candidates would accept the possibility of attending MR in other locations, even without this being the first option considered. One of these advantages is the possibility of immediate continuity of one's formation; another is the opportunity to experience other communities and cultures. This procedure is being followed since 2009 for targeting applicants from institutions of higher education participating in ENEM.¹⁴

Scholars in the field of medical education indicate the need for changes in the form of access to MR vacancies, especially with regard to objective proof of knowledge in the basic areas.^{2,15} This test generally has its items analyzed according to the Classical Test Theory (CTT), but this analysis is only possible after the application of the tests. Any problems with items are detected late, and often the quality of the selection is impaired. The problems with the items may occur in relation to the preparation, with conceptual problems, regarding discrimination or difficulty.

Considering the complexity involved in preparing a selection exam and that every institution in the various regions of the country is responsible for their selection processes to MR, it is easy to see that some of these tests have questionable quality both in relation to the instrument applied and the contents required. Thus, a unified examination could contribute to an improvement in the quality of selection, with its constructed instruments according to current concepts of validity and reliability, and could foster a retroactive effect for a uniform curriculum of undergraduate courses in Medicine in Brazil.

Entrance exams to medical residency programs

The selection process for MR, at the institution's discretion, may be conducted in two phases: written and practical. The first phase is mandatory and consists of a written, objective examination, with an equal number of questions in the specialties of Internal Medicine, Surgery, Pediatrics, Obstetrics and Gynecology, and Preventive and Social Medicine, weighing at least 50%. The second phase is optional, at the institution's discretion, consists of a practical test with a weight of 40-50% of the total score and curriculum analysis with maximum weight of 10% of the total grade.¹⁶

Currently, in Brazil, 53 specialties and 54 areas are recognized. Among the 53 specialties, 29 require direct access, i.e., the only prerequisite is a graduation degree in Medicine. For other specialties, requiring indirect access, prior training in the basic areas of knowledge is required.²

The proposed unification of processes would include the objective test in the first phase for specialties with direct

access, since currently these specialties are a priority for SUS. The other steps, with fewer competitors, could be decided by the institution.

Methodology

The proposed unified selection in medical residency programs

The selection processes for medical residency are admittedly necessary and have the pretention of differentiating students who have knowledge and skills considered important, in addition to being responsible for guiding the curriculum of undergraduate courses.

Although these tests satisfactorily fulfill a role in ranking the candidates, they have drawbacks. One of them is the decentralization of selection processes, which limitates the options in the race for one of the positions offered. On the other hand, they restrict the ability to recruit by institutions, disadvantaging those situated in locations further away from the major city centers.

One alternative change in the MR selection would then be the unification of vacancies through one single test, democratizing participation in selection processes and guiding young people in their fields of interest to underserved regions, adding to the current public policies for medical training and guidance of these professionals to priority areas for SUS. Furthermore, this unification of selection processes would be important, due to the possibility of promoting a curriculum integration among the various undergraduate courses in the country.

Use of item response theory in building the item bank

A high scale assessment as proposed in this study has a strong influence on educational policies, curricula of the institutions that are part of the process, and the professional future of young candidates, highlighting the importance of examining the many variables involved in the construction, application and scoring of the evaluation.

High scale evaluations commonly make use of a bank of items, which can be defined as a database of items relating some elements of each item, such as parameters, contents assessed, and others.¹⁷

Item parameters can be estimated by CTT, but it has some limitations. The main limitation is that measurements depend on the measurement instrument used (test) and on those assessed. This type of instrument can lead to errors and uncertainties in the selection process of candidates, since contests in general fill vacancies based on the raw scores obtained by the candidates in tests, being simply the count of questions answered correctly.^{18,19} This means that the score achieved by a candidate does not only depend on their own ability, but also on the level of difficulty of the test. Therefore, the comparison between responders can only be made if they undergo the same instrument.^{17,20}

An alternative to the currently widely used CTT in preparing high scale assessment is the Item Response Theory (IRT), which

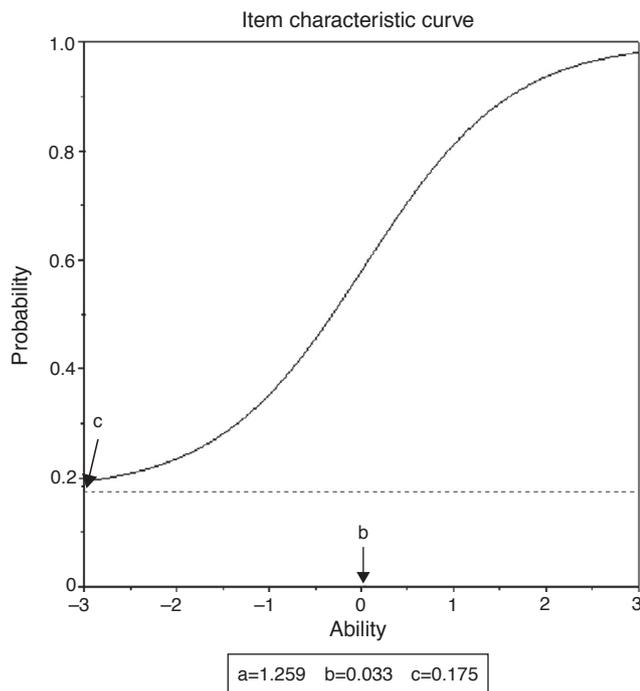


Fig. 1 – Characteristic curve of the item.

Source: BILOG-MG.²¹

consists of mathematical models to represent the relationships between the individual characteristics of the respondent (skills), and item characteristics (difficulty, discrimination, correct answer at random).

In IRT, the relationship between the probability of an individual answering correctly to an item and the parameters of this item is an incremental function, called item characteristic curve (ICC). Figure 1 illustrates the characteristic curve of an item considered efficient.²¹

Parameter b is a measure of item difficulty and is given in the same unit of skills (horizontal axis). In Figure 1, we notice that the higher the ability of the candidate, the greater his probability (vertical axis) to answer correctly to the item. The scale used is (0,1), i.e., the mean is zero and the standard deviation is 1. Drawing a vertical line on a skill, at the intersection of this line with CCI the probability of a candidate with that ability to answer correctly to the item is provided. Note that the further to the right is the ICC, the more difficult the item is.

IRT was developed in the United States in 1952 by Lord,²² and in the Netherlands in 1960 by Rasch.²³ This method became known mainly from 1968, with the work of Lord and Novick entitled *Statistical Theories of Mental Tests Scores*.²⁴

Some current examples of large-scale assessment using IRT in their item banks are *Graduate Record Examination (GRE)*, *Scholastic Assessment Test (SAT)*, *Test of English as a Foreign Language*²² (TOEFL), *Programa Internacional de Avaliação de Alunos*²⁵ (PISA), *Prova Brasil*,²⁶ *Sistema de Avaliação da Educação Básica*²⁷ (SAEB), ENEM,¹⁴ among others.

Using the templates provided by IRT, the items are designed to assess each latent trait and, from these, one can obtain the description of each item with some of its parameters, such as latent trait assessed, discrimination index, difficulty level, probability of success to chance. In IRT, the process of estimation of the item parameters is known as calibration.²⁰

Some advantages of the unification of the selection process in MR programs by means of a calibrated item bank according to the IRT include:

- Using pre-tested items: items are classified according to difficulty, discrimination, probability of success at random, among others. Faulty items in preparation are discarded.
- Developing a range of unique skills: item parameters and abilities of respondents are placed on a common scale, making it possible to compare candidates and item parameters, even if from distinct tests or groups.
- Allowing student mobility: enables candidates to participate in the selection process in their city of origin, without the need to travel to the location of the desired institution. This is also a way of democratizing access, not prioritizing the local egress over candidates from outside.
- Developing several editions of selection test: equivalent tests are organized as needed. Candidates are graded on a single scale, so idle vacancies can be avoided, even in the institutions of the least favored regions.
- Deepening analyses: a qualitative analysis of the items and the respondent population is possible, from the answers of individuals.
- Guiding education: standardization of curricula of all undergraduate medical courses in the country.

Professional associations, the government and scholars highlight the need for changes to the MR and particularly the access way, as noted earlier in this article. It is expected, therefore, that the creation of a bank of items to be applied to unified selection exams for MR, calibrated according to item response theory, provides a substantial gain in quality to current processes of selection, meeting the demand of institutions for a more efficient selection and also the needs of SUS as for redistribution of professionals in the country.

Results

Currently, the selection tests for MR programs are developed and implemented in a decentralized way, each institution is responsible for the assessment. Experts question the quality of these tests, their subjectivity and lack of fairness, since many of these places are disputed by both local candidates and by candidates coming from other city centers.^{2,12}

The institutions responsible for MR programs not usually leave on their official websites reports with studies about the quality of the items in their tests, or about the validity and reliability of their instruments. In fact, in Brazil, the number of researches on high scale assessments can still be considered limited. Researchers of the assessment field criticize both the scarcity of studies on the issue of Brazilian tests, discarded without assessing their validity, reliability and impacts

in education and society, and studies on the poor official disclosure of the results of these examinations by the agencies responsible.^{28,29}

This scarcity of studies on high scale assessments in Brazil results in poor transparency of the processes involved in the preparation, correction and scoring of exams. Another factor also related to this poor literature is the lack of resources and other support from the government and agencies responsible, able to support and encourage research on these assessments.

Problems with the quality and difficulty of items are common and can be detected only after using the instrument.^{18-20,30,31} This is the reason why it requires pre-tests in building an item bank. In fact, when developing an item bank for assessing learning contents in Spain, Garcia et al.³² excluded about 25% of the items, a number similar to that achieved when building the bank Hezinet, a multimedia system used to Basque language teaching to foreigners.³⁰ In its methodology to develop the item bank and constant selection of new items to be included in this bank, ENEM uses pre-testing in a systematic way to high school students across the country.²⁶

Discussion

Problems with items in high scale tests, such as vestibular (an university entrance exam in Brazil), public tenders or selection examinations for MR are common, but few are reported in the press or on the websites of the institutions responsible. Candidates are only informed of items canceled due to conceptual problems. Problems with the quality of items, such as discrimination, difficulty or probability of right answer at random are not normally disclosed. It is also not common to disseminate studies on the validity and reliability of the instruments used in these exams.

Studies prove that 20-25% of items of an instrument are normally excluded, due to conceptual problems or inefficiency as for discrimination. An item bank developed using correct pedagogical procedures, pre-tested before composing the assessment instrument, has a great advantage, and with quality often superior to traditional tests that are part of the selection programs for MR.

To elaborate a good assessment instrument, the items should be assessed in pedagogical, quantitative and qualitative terms. These analyzes must verify whether the assessment meets the objectives and criteria for validity and reliability. IRT models allow greater flexibility in the drafting process of tests and in analyzing the answers given by the people assessed, in addition to resulting in a more fair and consistent classification.^{18-20,31}

The entire procedure of drafting items and pre-testing to compose a large item bank calibrated as per IRT is set and well-defined since the reformulation of ENEM in 2009.¹⁴ ENEM now occupies a prominent place in the Brazilian education sector for its contribution in the democratization of access to higher education, unifying this access to a large number of universities.²⁶

The knowledge acquired through the ENEM experience, if replicated in the unified selection for MR programs, could provide a substantial gain in quality to current processes

of selection, meeting the demand of institutions for a more efficient selection and also the needs of SUS as for redistribution of professionals in the country.

Conclusion

One of the problems faced by medical education is in expanding the number of courses without minimum guarantees on the quality of the course.^{1,3} The educational quality requires systematic assessments and pre-set targets. Assessment is important in any education system, causing changes needed and monitoring teaching and learning.^{28,33,34}

Furthermore, the proposed unification of the selection examination for MR could also serve to assess the under-graduation courses in Medicine. Also, due to the possibility of IRT models to compare different tests, and also compare individuals who answer the tests, these assessments could provide data for many studies, including longitudinal.

Another problem highlighted in the specific literature is the misdistribution of specialized physicians in the various regions of the country. Physicians play an essential role for the health care of the population, but the country faces serious problems regarding lack of assistance in some isolated areas of large cities.^{1,5,7,13} Political actions should be directed to the training of qualified specialists, according to the health needs of each region of the country. The proposed changes in selection examinations of MR are always mentioned by experts as one of the ways to deal with this problem.^{7,10,11}

As in ENEM,¹⁴ which aims to select candidates for positions in under-graduation courses, the use of a unified examination for MR programs will allow candidates to compete for vacancies offered by all institutions, thus promoting regional mobility.

Conflicts of interest

The authors declare no conflicts of interest.

REFERENCES

1. Michel JLM, Oliveira RAB, Nunes MPT. Residência médica no Brasil. In: Cadernos da ABEM; 2011. v.7. Rio de Janeiro: Associação Brasileira de Educação Médica; 2011.
2. Nunes MPT, Michel JLM, Haddad E, Brenelli SL, Haddad DM, Ribeiro ECO, Petta HL. Distribuição de vagas de residência médica e de médicos nas regiões do país. In: Cadernos da ABEM, v. 7, out. 2011. Rio de Janeiro: Associação Brasileira de Educação Médica; 2011.
3. Massuda A, Cunha FM, Petta H. Residência médica: contribuições dos médicos residentes ao debate. *Rev Assoc Med Bras.* 2007;53(2):96-7.
4. Nunes MPT. Os 3 grandes desafios da Residência Médica. CREMESP: Conselho Regional de Medicina do Estado de São Paulo; 2010. ed. 53.
5. Petta HL. Formação de médicos especialistas no SUS: descrição e análise da Implementação do Programa Nacional de Apoio à Formação de Médicos Especialistas em Áreas Estratégicas

- (Pro-residência) [dissertation]. Rio de Janeiro: Escola Nacional de Saúde Pública Sérgio Arouca –IPEA; 2011.
6. Avery DM Jr, Wheat JR, Leeper JD, Mcknight JT, Ballard BG, Chen J. Admission Factors Predicting Family Medicine Specialty Choice: A Literature Review and Exploratory Study Among Students in the Rural Medical Scholars Program. *J Rural Health*. 2012;28(1):128-36.
 7. Campos FE, Machado MH, Girard NS. A fixação de profissionais de saúde em regiões de necessidades. *Divulg Saúde Debate*. 2009;(44):13-24.
 8. Leão PBOS, Martins LAN, Menezes PR, Bellodi PL. Well-being and help-seeking: an exploratory study among final-year medical students. *Rev Assoc Med Bras*. 2011;57(4):379-86.
 9. Lourenção LG, Moscardini AC, Soler ZASG. Saúde e qualidade de vida de médicos residentes. *Rev Assoc Med Bras*. 2010;56(1): 81-91.
 10. Silva RB, Pinealt R. Impact of physician distribution policies on primary care practices in rural Quebec. *Can J Rural Med*. 2012;17(3):92-8.
 11. Ximenes JAF, Silva MGC. Residência médica em otorrinolaringologia no Ceará em 2003: oferta de vagas e perfil da concorrência. *Rev Bras Otorrinolaringol* 2006;72(6):826-30.
 12. Propostas aprovadas no XII Enem: Formação Médica – Residência médica. XII Encontro Nacional de Entidades Médicas. Brasília (DF), 28 a 30 jul. 2010 [accessed 23 Set 2012]. Available at: <http://medico.cfm.org.br/enem/index.php/pre-enem/29-tema-i-formacao-medica/69-residenciamedica>
 13. Corrêa NA. Caracterização da procura por especialidades médicas dos candidatos ao Concurso SUS – SES/SP entre 1999 e 2004 [dissertation]. São Paulo: Faculdade de Medicina, Universidade de São Paulo; 2008.
 14. Exame Nacional do Ensino Médio [accessed 18 Aug 2010]. Available at: <http://www.enem.inep.gov.br/>
 15. Ribeiro MAA. Apontamentos sobre Residência Médica no Brasil. Brasília (DF): Câmara dos Deputados – Consultoria Legislativa; 2011. Biblioteca Digital Câmara. [accessed 23 Set 2012]. Available at: <http://bd.camara.gov.br>
 16. Secretaria de Educação Superior. Comissão Nacional de Residência Médica. Resolução n. 3, 16 set. 2011.
 17. Moreira Junior FJ. Sistemática para implantação de testes adaptativos informatizados baseados na teoria da resposta ao item [thesis]. Florianópolis: Faculdade de Engenharia Produção; Universidade Federal de Santa Catarina; 2011.
 18. Andriola WB. Doze motivos favoráveis à adoção do Exame Nacional do Ensino Médio (ENEM) pelas Instituições Federais de Ensino Superior (IFES). Ensaio: Aval Pol Públ Educ. 2011; 19(70):107-25.
 19. De Ayala RJ. The theory and practice of item response theory. New York: The Guilford Press; 2009.
 20. Andrade DF, Tavares HR, Valle RC. Teoria da resposta ao item: conceitos e aplicações. SINAPE; 2000.
 21. Zimowski MF, Muraki E, Mislevy RJ, Bock RD. Bilog-MG. Lincolnwood: Scientific Software Inter-national; 2002.
 22. Lord FMA. A theory of test scores. Psychometric Monograph; 1952.
 23. Rasch G. Probabilistic models for some intelligence and attainment tests. Copenhagen: Danish Institute for Educational Research; 1960.
 24. Lord FMA, Novick NR. Statistical Theories of mental test scores. Massachusetts: Addison Wesley; 1968.
 25. Programa Internacional de Avaliação de Alunos (PISA) [accessed 23 Set 2012]. Available at: <http://www.enem.inep.gov.br/internacional>
 26. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira – INEP [accessed 22 Aug 2012]. Available at: <http://www.inep.gov.br>
 27. Sistema Nacional de Avaliação da Educação Básica – SAEB [accessed 1 Set 2012]. Available at: <http://www.inep.gov.br/saeb/>
 28. Scaramucci, MVR. Efeito retroativo da avaliação no ensino/aprendizagem de línguas: o estado da arte. *Trab Ling Aplic*. Campinas. 2004;43(2):203-26.
 29. Vianna HM. Avaliações nacionais em larga escala: análises e propostas. *Estudos em Avaliação Educacional*, n. 27, jan-jun; 2003.
 30. Cuadrado JL, Pérez TA, Vadillo J A, Gutiérrez J. Calibration of an item bank for the assessment of Basque language knowledge. *Comput Educ*. 2010;55(3):1044-55.
 31. Embretson SE, Reise SP. Item response theory for psychologist. Mahwah: Lawrence Erlbaum Associates; 2000.
 32. García EGJ, Rodríguez G. Elaboración de un TAI sobre contenidos escolares. In: Olea J, Ponsoda V, Prieto G. *Tests informatizados: Fundamentos y aplicaciones*. Madrid: Pirámide; 1999.
 33. Nunes MPT, Martins ACS, Machado V R. Avaliação do Sistema RM: instituição, programa, residente, corpo docente. In: *Cadernos da ABEM*, v. 7, out. 2011. Rio de Janeiro: Associação Brasileira de Educação Médica; 2011.
 34. SisCNRM: Sistema da Comissão Nacional de Residência médica [accessed 14 Set 2012]. Available at: http://portal.mec.gov.br/index.php?id=12233&option=com_content&task=view