Author's response to reviews

Title: Language, culture and international exchange of virtual patients.

Authors:

Valentin Muntean (valentin.muntean@gmail.com)
Tudor Calinici (tcalinici@umfcluj.ro)
Stefan Tigan (stigan@umfcluj.ro)
Uno GH Fors (Uno@dsv.su.se)

Version: 2 Date: 19 December 2012

Author's response to reviews: see over
Cover letter to Editorial Board, BMC Medical Education

Dear Editors,

We provided a point-by-point response to the reviewer concerns and made the asked revisions of the manuscript. We also answered the Editorial Requests.

With best wishes,
Valentin Muntean
Tel: 00 40 722 704 401
e-mail: valentin.muntean@gmail.com

Reviewer 1

Mr. Fabrizio Consorti

Major compulsory revisions

1. Abstract: the abstract is in some way misleading when it states "The students worked on eight VPs". Each student worked only on two VPs out of 4+4. The rest of the abstract does not clarify this point, which is methodologically crucial.

The misleading statement was changed to: (Abstract, Method)

The students worked on eight VPs in two identical versions, Romanian and English. The first group (2010) of 136 students worked with four VPs developed in Cluj and the second group (2011) of 144 students with four VPs originally developed at an US University. Every student was randomly assigned two different VPs, one in Romanian and another in English.

2. page 3, end of Background: "The present research tried to unequivocally answer the question before mentioned". There is not any question before this sentence. Since this study aims to be a quasi experimental study, please state in an explicit way the research hypothesis

The research hypothesis was stated explicitly: (Background, 3rd paragraph)

The present research tried to find out if using VP’s in English would be a viable option for developing a VPs program in our institution.
3. the study considers two distinct groups of students of two different years. In the Discussion (pg. 8 line 19) the authors state "The 2010 and 2011 groups of students had similar background and training", but since a pre-test of any kind was not performed, there is no evidence that the two groups had a similar level of knowledge and skill. Please, acknowledge this point in your Discussion.

We acknowledged that, since a pre-test was not performed, there is no direct evidence of similar knowledge and skill of the two groups. (Discussion, 4th paragraph)

Even if a pre-test was not performed, we presumed that the two groups had a similar level of knowledge and skill.

4. Study design: the authors should provide justification of their choice of submitting only two (1 Romanian + 1 English) out of 4 possible VPs. (The choice is among 4, because presumably the same case was not given to a student in both the Romanian and English version). Why did the author choose to halve the number of VPs for each student, in considering that the higher the number of items in a test, the higher is the reliability of the test? This is even more relevant in dealing with clinical cases, where the "case specificity effect" described by G.Norman is at work.

We decided to give only two different VPs to each student, one in Romanian and one in English. The mean time per case was 42-47 minutes and the students had to do the two cases during a seven day period. That corresponds to the two hours allocated per week for the independent activity in the optional course “Methods of teaching and evaluation for medical students”. Moreover, our research focused on comparison of student group performance on the Romanian versus English version of the cases rather than individual student performance, which would need far more than two - four cases for an acceptable reliability.

Explanation added to the Methods: (Methods, 3rd paragraph)

The cases were supposed to be solved in the two hours allocated per week for the independent activity in the optional course “Methods of teaching and evaluation for medical students”.

5. Checklist: “The diagnostic, therapeutic plan and diagnosis justification were analyzed and graded using checklists developed by the case authors”. It seems that different case authors developed different checklists, but this fact could strongly hamper the reliability of the assessment. Were the checklist validated in any way? Were guidelines to develop the checklists in a consistent way available?

The VPs and checklists were developed by the teachers from the Internal Medicine Department of Cluj-Napoca. For the two identical versions, Romanian and English, of each VP, the same checklist, developed by the case author, has been used. Cases and checklists were developed based on practice guidelines for diagnosis and treatment provided by the Internal Medicine Department of Cluj-Napoca. Every case was peer reviewed in terms of content, design and media by two teachers from Commission for Students. The
cases used in the study have been used for student training and evaluation for more than two years and rated in student evaluation as “very good” and “of medium difficulty”.

Explanation added to the Methods: (Methods, 2nd paragraph)

For the two identical versions, Romanian and English, of each VP the same checklist, developed by the case author, has been used. Cases and checklists were developed based on practice guidelines for diagnostic and therapy provided by the Internal Medicine Department of Cluj-Napoca. Every case was peer reviewed in terms of content, design and media by two teachers from Commission for Students. The cases have been used for student training and evaluation for more than two years and rated in student evaluation as “very good” and “of medium difficulty”.

Explanation added to the Methods: (Methods, 5th paragraph)

The second group of students (2011) worked on four VPs originally developed in the USA, with the Web-SP software –Fig.1 and 2. In student evaluation the cases were of medium difficulty and very good, in terms of content, design and media. The translation into Romanian was made by teachers from the Internal Medicine Department. For the two identical versions, English and Romanian, of each VP, the same checklist, developed by the case translator, has been used. The translation into Romanian was made by teachers from the Internal Medicine Department. For the two identical versions, English and Romanian, of each VP the same checklist, developed by the case translator, has been used.

6. Interactions: why were only the first 25 interactions considered? what does it mean? Twenty-five for each section? what about a student not reaching the end of the exercise within 25 interactions? How can the mean number of history questions be 49 for example (see Table 1) if you cut the interactions to the first 25? This point is not clear, please explain better and in more details.

The 131 remaining students in the first study finalized 262 interactions, half in Romanian and half in English. The students were randomly assigned two different cases, one in English and one in Romanian. Because of randomization of cases and the exclusion form the study of the five students who did not complete both cases, in the end the number of interactions available for the 8 VPs varied between 25 and 40. For the 140 students in the 2011 group the (280 cases finalized) the number of interactions/VP varied between 26 and 44. In order to limit the influence of the inter-case variability on group performance on the cases we decided to select for statistical analysis the same number of interaction for each VPs. For each of the 8 VPs versions (4 VPs in Romanian and 4 VPs in English), we selected for the statistical analysis, the chronologically first 25 interactions (history questions, exams, lab/imaging tests etc.) per case (8x25=200 interactions). The variability of the mean number of questions for history, physical examination and laboratory is a characteristic of Web-based Simulation of Patients (Web-SP) system software, which allows the student to ask hundreds of different questions.

Explanation added to Methods: (Methods, 4th and 6th paragraphs)
Because of randomization of cases and the exclusion form the study of the five students who did not complete both cases, in the end the number of interactions available for the 8 VPs varied between 25 and 40.

Because of the randomization of cases and the exclusion from the study of the four students who did not complete both cases, in the end the number of interactions available for the 8 VPs varied between 26 and 44.

7. Tables:

- Table 1 and Table 2 seem to be not aligned. For example in Table 1 the mean time/case for VPs developed in Cluj in 2010 in Romanian is 47.00 s.d. 25.69. In table 2 the same variable is 47 s.d. 31 (row 4 of the table). Please check carefully the consistency.

All inconsistency in between the two tables has been cleared.

In the second table the figures were derived from an initial analysis, in which we included all interactions (262 respectively 280 interactions). In both tables there are now figured the first 25 interactions for each of the 8 VPs in 2010 and 8 VPs in 2011 (200 respectively 200). Thank you for observing that.

- Why did the author use a one-tailed test in Table 1? Was the experimental hypothesis strong enough to allow this assumption?

In table 1, we replaced now the one tailed test result with the two tailed results. However, for two-tailed test the differences between English and Romanian are not significant but for one tailed test there exist one significant difference.

- Please explain what does Asymp.Sig. in Table 2 mean and how does it refer to the Mann-Whitney test quoted in Methods

Asymp.Sig. in Table 2 means p-value for Mann-Whitney test / changes done in Table 2.

Minor Essential revisions

line 11 in Methods: "[cases]... were considered of medium difficulty and very good, in terms of content, design and media." Who assessed the cases? Where they self-assessed by the authors or were they assessed by external reviewers?

Please specify

Explanation added to the Methods: (Methods, 2nd paragraph)
Every case was peer reviewed in terms of content, design and media by two teachers from Commission for Students. The cases have been used for student training and evaluation for more than two years and rated in student evaluation as “very good” and “of medium difficulty”.

line 15 in Methods: "The total repurposing time for the four cases". The beginning of the paragraph just states "VPs developed by the academic staff" without any number, so the reader must guess that they were four. Please specify

The requested specification was added to Methods: (Methods, 2\textsuperscript{nd} paragraph)

Four cases were initially developed in Romanian, and then translated into English by the author. The content and details of the versions were identical, and the translation into English was reviewed by a native English speaker. The total repurposing time for the four cases (translation into English, editing, review and final check) was 30, 25, 33 and respectively 36 hours (mean 31 hours).

line 8 pg. 4: "free text questions re. diagnosis, therapy". What does re. mean?

Re. is an abbreviation for regard.

The abbreviation has been removed. Methods, 3\textsuperscript{rd} paragraph

The students were given no recommendation on which language, Romanian or English, they should use for answering the free text questions regarding diagnosis, therapy and justifications.

Quality of written English: Not suitable for publication unless extensively edited

The written English was reviewed by a native English speaker, Mrs. Sally Wood-Lamont, Editor and Reviewer at the “Valeriu Bologa” Library, “Iuliu Hatieganu” University of Medicine and Pharmacy, Cluj-Napoca, Romania. http://www.umfcluj.ro/Detaliu.aspx?t=Biblioteca-Date-de-contact

Reviewer 2

Mr. Afonso Cavaco

Discretionary Revisions.

This is a research paper where it is easy to find a research objective, although it was not formulated as a clear research question.

The research hypothesis was stated explicitly: (Background, 3\textsuperscript{rd} paragraph)

The present research tried to find out if using VP’s in English would be a viable option for developing a VPs program in our institution.
The methods are appropriate and well described. Anyway, some reviewer’s comments are presented below.

The data collection and the analytical methods used are sound, being the discussion and conclusions balanced and supported by the data. Nevertheless, based on the manuscript’s title, some discussion on cultural variations of VP dialogues could have been more explored along the manuscript.

(Discussion, 11th paragraph)  
Besides the specific content, design and language, VPs include also “cultural” features [6], which give a specific flavor and authenticity of the case. Those are difficult or impossible to translate. Moreover, VPs can be used for learning the background and the medical conditions of patients from different countries and cultures [6]. In our research we found many statistically significant differences in student activity through the cases, between those with Cluj authors (2010 group) and those developed in the USA (2011 group). These differences persist when we separately compared the Romanian and English versions of the two sets of VPs, so it can be deduced that they are not related to the language proficiency of students and are probably related to the “cultural” features of the cases, difficult or impossible to translate.

No limitations are stated by authors, although some comments might be made in relation to cultural adaptation and the level of virtual reality that was being used.

Methods: (Methods, 2nd paragraph)  
The software used was the Web-based Simulation of Patients (Web-SP) system, originating from the Karolinska Institutet, Stockholm, Sweden [11, 12].

(Discussion, 3rd paragraph)  
The design of the cases edited in the Web-based Simulation of Patients (Web-SP) was ‘narrative’, not only best suited for history-taking and communication skills [16], including diagnostic ability [4], but also for learning and the assessment of clinical reasoning [5, 17, 18].

Explanation added to the Methods: (Methods, 5th paragraph)  
The second group of students (2011) worked on four VPs originally developed in the USA, with the Web-SP software –Fig.1 and 2. In student evaluation the cases were of medium difficulty and very good, in
terms of content, design and media. The translation into Romanian was made by teachers from the Internal Medicine Department. For the two identical versions, English and Romanian, of each VP, the same checklist, developed by the case translator, has been used. The translation into Romanian was made by teachers from the Internal Medicine Department. For the two identical versions, English and Romanian, of each VP the same checklist, developed by the case translator, has been used.

Without questioning the true merits of the ‘eViP’ initiative and knowing the intense resources consumption of most VPs solutions, authors have not discussed how much this web-based simulation of patients (Web-SP) differs from common computer case-study presentations i.e. how much the teaching/learning process would benefit from other computer solutions, including different levels of immersive virtual reality.

In ‘Iuliu Hatieganu’ University of Medicine the VPs are used in PBL course during the preclinical years and in “blended learning” during clinical years.

Discussion, 2nd paragraph
As partners in the three-year EU-funded project, called ‘eViP, we started to repurpose some Cluj Faculty of Medicine VPs in the Web-based Simulation of Patients (Web-SP) system from Sweden.

Discussion, 3rd paragraph
The design of the cases edited in the Web-based Simulation of Patients (Web-SP) was 'narrative', not only best suited for history-taking and communication skills [16], including diagnostic ability [4], but also for learning and the assessment of clinical reasoning [5, 17, 18]. The criteria used for the assessment of the students’ activity on the cases were those offered by the software, namely ordering correct illness history, physical exam, lab/imaging tests, and suggesting correct diagnosis, therapy and justifications of those. The students’ ability to obtain the correct diagnosis and treatment was assessed and graded using checklists developed by the case authors.

Additionally, the study here reported is also missing feedback from students, their experience with the system (being first users), including VPs acceptability and perceived value as a pedagogical option for learning and/or assessment in medical education.

Discussion, 2nd paragraph
The VPs were initially used in PBL course during the first two preclinical years and later, in clinical years, as "blended learning". The feedback received from students was positive. Student evaluation rated VPs funny, engaging and informative, with great value in learning and evaluation.

Authors have referred previous published work done in this area and there are no comments in relation to the abstract, except for the last sentence of the methods section, which is not clear to me.
Good knowledge of English is mandatory for all our students and printed and electronic materials in English, French or German, as recommended readings for lectures, seminars and clerkships are regularly used. That is why, three years ago, when we started with the VPs edited in the Web-SP software during the optional course “Methods of teaching and evaluation for medical students”, VPs in both Romanian and English were utilised. The initial impression was that Romanian students’ performance on Romanian VPs is better than on the English versions (better diagnosis and treatment plan) [6]. Fors UG, Muntean V, Botezatu M, Zary N: Cross-cultural use and development of virtual patients. Med Teach 2009, 31(8):732-738.

Minor Essential Revisions.

The last sentence of the last paragraph in the Background section refers to a question. Since no questions were formulated previously, I would suggest replacing this word by “issue”.

The “question” was removed: (Background, 3rd paragraph)

The present research tried to find out if using VP’s in English would be a viable option for developing a VPs program in our institution.

In the Methods section (3rd paragraph) authors refer students were randomly assigned with two different cases (English and Romanian). There are no descriptions to what each case comprised i.e. how much similitude/dissimilitude existed between VPs, thus controlling for possible variation between clinical situations.

The specification was added to Methods: (Methods, 2nd paragraph)

Four cases were initially developed in Romanian, and then translated into English by the author. The content and details of the versions were identical, and the translation into English was reviewed by a native English speaker. For the two identical versions, Romanian and English, of each VP, the same checklist, developed by the case author, has been used. Cases and checklists were developed based on practice guidelines for diagnostic and therapy provided by the Internal Medicine Department of Cluj-Napoca. Every case was peer reviewed in terms of content, design and media by two teachers from Commission for Students. The cases have been used for student training and evaluation for more than two years and rated in student evaluation as “very good” and “of medium difficulty”.

The specification was added to Methods: (Methods, 5th paragraph)

The second group of students (2011) worked on four VPs originally developed in the USA, with the Web-SP software –Fig.1 and 2. In student evaluation the cases were of medium difficulty and very good, in terms of content, design and media. The translation into Romanian was made by teachers from the Internal Medicine Department. For the two identical versions, English and Romanian, of each VP, the
same checklist, developed by the case translator, has been used. The content of the versions was identical, and the translation was reviewed by a native English speaker.

Major Compulsory Revisions.

In the Results section (2nd paragraph), authors have found a statistically significant difference in the number of words in diagnostic justification. Since Romanian is a Latin-based language differing from English (Germanic-based), it is expected some natural variation, with a probable greater number of words in Romanian. This raises doubts in relation to the true meaning of differences found in the number of words used in both VPs languages.

**Results, 2**\textsuperscript{nd} and 3\textsuperscript{rd} paragraphs:

For 2010 group student answers, diagnosis, therapeutic plan and diagnosis justification, were in Romanian 171 times and in English 29 times, all for VPs in English. For 2011 group most student answers were in Romanian, 174 times, and only 26 times in English, all for VPs in English.

That makes me think that language influence on number of words used in diagnosis justification is without major influence on final result.

Additionally, authors have also registered students’ preference for answering in Romanian rather than English, when using the English VP. Therefore, the last column of Table 1 should reflect this situation and discriminate between students’ replies, for English VPs.

In the last column in Table 1 RO stands for the Romanian version of VPs and ENG for the English version of the same cases. For both, most of the answers were given in Romanian language (for all Romanian versions of VPs and for most of the English versions of VPs, but 29+26). The small number of answers in English would make difficult any comparison with answers in Romanian...

In Results 3rd paragraph, authors described to have found a number of statistically significant differences in students activities between the 2010 and 2011 cohorts. Reasons supporting this are discussed later in the Discussion section, 4th paragraph. Using an epidemiological perspective, it is hard to accept that both cohorts had an equivalent background and training without giving any evidence-based data. In paired comparisons such as this one, researchers should have controlled for possible background differences and confounding effects. Authors should have demonstrated how the 2 cohorts were similar in basic demographics (e.g. average age and gender) as well as in education performance (e.g. average marks). Otherwise, the differences found might not be due to VPs origin (Cluj vs. USA). Unwanted bias might have been introduced if both 2010 and 2011 groups were already different in their learning process and academic success.

The Discussion 4\textsuperscript{th} paragraph has been changed to:

Even if a pre-test was not performed, we presumed that the two groups had a similar level of knowledge and skill. The 2010 and 2011 groups of students had similar background and training. All fourth year medical students participating in the research had previously studied Internal Medicine and none had
worked with VPs before. The previous training of the students was exclusively in Romanian and all of them had good English proficiency (equal to or better than the B1 level).

There were no differences of age and gender between the two groups of students. Actually, it would extremely difficult or impossible to control education background and academic performance of two students groups in “real academic life”, i.e. non-experimental setting.

Methods, 3rd paragraph
Access to the cases was given for a seven day period and the students logged on through the Internet in their homes. The cases were supposed to be solved in the two hours allocated per week for the independent activity in the optional course “Methods of teaching and evaluation for medical students “.

Moreover, the number of clinical cases necessary to evaluate student performance with a good reliability should be more than 20, again very difficult or impossible to reach having in mind the “actual curricular constraints”. The present research focused on our students from Faculty of Medicine and actual curricular and academic environment, to find out if using VP’s in English would be a viable option for developing a VPs program in our institution.

I also found difficult to understand the end of the 2nd sentence in this 3rd paragraph.

The necessary correction was made: (Results, 3rd paragraph)

The 2010 group of students spent more time on the cases, asked for a lesser number of history questions and physical queries, and more laboratory and imagistic information, and used less words in diagnosis justification when compared with 2011 group.

As referred previously, the 4th paragraph in the Discussion section would benefit from more evidence in Results. This is further justified when in the 11th paragraph authors discussed differences having isolated the language effect. In this 11th paragraph is also stated that cultural features are difficult to translate into VPs, suggesting that cultural equivalency was not possible to address when, for instance, USA VPs were translated to Romanian. Although this confounder is indeed hard to control for, here authors had an interesting opportunity to discuss study limitations. This is even more relevant since language and culture are part of the manuscript title.

The 11th paragraph was changed to:

Besides the specific content, design and language, VPs include also “cultural" features [6], which give a specific flavor and authenticity of the case. Those are difficult or impossible to translate. Moreover, VPs can be used for learning the background and the medical conditions of patients from different countries and cultures [6]. In our research we found many statistically significant differences in student activity through the cases, between those with Cluj authors (2010 group) and those developed in the USA (2011 group). These differences persist when we separately compared the Romanian and English versions of the two sets of VPs, so it can be deduced that they are not related to the language proficiency of students and are probably related to the “cultural" features of the cases, difficult or impossible to
translate. We found no statistical differences the ability of the students (2010 group versus 2011 group) to reach the correct diagnostic and therapeutic plan.

VPs can be used for learning the background and the medical conditions of patients from different countries and facilitate the mobility of healthcare professionals and globalization in medical education. The VPs in native language of the patient offer the “cultural” features which give a specific flavor and authenticity. The conclusion of our study is that the development of medical education programs with VPs in English would be feasible, cost-effective, and in accordance with the globalization in medical education.

Quality of written English: Acceptable

The written English was reviewed by a native English speaker, Mrs. Sally Wood-Lamont, Editor and Reviewer at the “Valeriu Bologa” Library, “Iuliu Hatieganu” University of Medicine and Pharmacy, Cluj-Napoca, Romania. http://www.umfcluj.ro/Detaliu.aspx?t=Biblioteca-Date-de-contact

Editorial Requests:

1. We recommend that you ask a native English speaking colleague to help you copyedit the paper. If this is not possible, you may need to use a professional language editing service. For authors who wish to have the language in their manuscript edited by a native-English speaker with scientific expertise, BioMed Central recommends Edanz (www.edanzediting.com/bmc1). BioMed Central has negotiated a 10% discount to the fee charged to BioMed Central authors by Edanz. Use of an editing service is neither a requirement nor a guarantee of acceptance for publication. For more information, see our FAQ on language editing services at http://www.biomedcentral.com/authors/authorfaq/editing.

The written English was reviewed by a native English speaker, Mrs. Sally Wood-Lamont, Editor and Reviewer at the “Valeriu Bologa“ Library, ”Iuliu Hatieganu“ University of Medicine and Pharmacy, Cluj-Napoca, Romania. http://www.umfcluj.ro/Detaliu.aspx?t=Biblioteca-Date-de-contact

2. Ethics - Experimental research that is reported in the manuscript must have been performed with the approval of an appropriate ethics committee. Research carried out on humans must be in compliance with the Helsinki Declaration (http://www.wma.net/en/30publications/10policies/b3/index.html), and any experimental research on animals must follow internationally recognized guidelines. A statement to this effect must appear in the Methods section of the manuscript, including the name of the body which gave approval, with a reference number where appropriate.

The research design, including ethical issues, was approved by the Dean, the Curriculum Office and the Ethics Committee of the Cluj-Napoca Faculty of Medicine - No.57/23.03.2010.
3. Please clarify if you obtained permission to publish the images of the individual pictured in figures of 1 and 2.

We change the pictures from Sweden / USA with pictures of our patients, from whom we have the written permission to publish.

Please also ensure that your revised manuscript conforms to the journal style (http://www.biomedcentral.com/info/ifora/medicine_journals). It is important that your files are correctly formatted.

The revised manuscript conforms to the journal style.