ASSESSING ASSESSMENT TOOLS: TOWARDS QUESTIONNAIRE IMPROVEMENT THROUGH VALIDATION

Maria João Fonseca
Universidade do Porto

Patrício Costa, Leonor Lencastre
Universidade do Porto

Fernando Tavares
Universidade do Porto

Abstract

A major concern of Science Education researchers has been to assess the effectiveness of educational strategies, namely curricular structure, alternative teaching approaches and informal learning environments. In these studies, parameters such as knowledge, attitudes and behaviors are frequently addressed using diverse instruments chosen to answer precise research questions. In order to adopt corrective measures it is fundamental to obtain reliable results. If case-study approaches can be suitable to study in detail specific teaching activities, they hardly give a wider view of the students' universe. In contrast, questionnaires are commonly used assessment tools that allow broad surveys of multiple aspects among a vast universe of students. Furthermore, they provide considerable amounts of easily treatable information in relatively short periods of time. For these reasons, educational observatories, as the well-known EUROBAROMETER, tend to favour inquiry-based assessment. Regardless of their apparent simplicity, questionnaire-based instruments require proper procedures to provide reliable and valid data. To comprehensively assess the knowledge, interest and attitudes of elementary and high-school students from Porto towards biotechnology, and identify the sources they use and trust to gather information about it, an oriented questionnaire validation guideline is proposed, combining pilot work with statistical validation through exploratory factor analysis and reliability analysis.

Introduction

The ultimate goal of science education is the promotion of scientific literacy (Braun & Moses, 2004; Cabo Hernandez, Mirón, & Cortiñas Jurado, 2006), which is essential to assure that knowledgeable citizens are competent to take part in the democratic decision-making policies regarding scientific and technological processes. Biotechnology is one of the most controversial socio-scientific issues to date (Dawson, 2007; Klop & Severiens, 2007). Acknowledging the range and intensity of its impact on society, many studies have been conducted addressing public knowledge and attitudes towards biotechnology (Ping & Gutteling, 2009). However, the ones targeting student populations are still scarce (Dawson & Soames, 2006; Dawson, 2007; Firman, 2007; Klop & Severiens, 2007; Prokop, Lešková, Kubitasko, & Diran, 2007; Sáez, Niño, & Carretero, 2008, Uşak, Erdoğan, Prokop, & Özcel, 2009).

Quantitative inquiry surveys are the most suitable methodologies to achieve a broad characterization of a given population, especially when planning to establish correlations and make generalizations (Oppenheim, 1992, Black 1999). Nevertheless, questionnaires have limitations that are mainly related to their subjectivity (Black, 1999). Thus, it is necessary to design and administer these instruments following proper procedures that can assure the validity and reliability of the results they provide (Brace, 2008). Although validation should be an intrinsic element of large scale assessment, many studies fail to report these procedures (Petric & Czad, 2003). The decision to dismiss
questionnaire validation may result from the perception that these statistical methods are not efficient given the time and effort they require, particularly since most research designs have strict timings (Donney, 2002).

In this context, we designed a questionnaire to make the diagnostic of the knowledge, attitudes and decision-making capacity regarding biotechnology of elementary and high school students engaged in two different curricular formats: with and without Biology contents. The consistency of this questionnaire was improved through a pilot study and using statistical approaches to increase the data’s reliability. Exploratory factor analysis and reliability analysis were conducted and their outcomes discussed.

**Rationale**

Questionnaire validation through a pilot study allows to identify, understand and address ambiguities that can interfere with the students’ answers and compromise the reliability of the data gathered (Oppenheim, 1992, Black, 1999). It is important to assess the adequacy of the questionnaire’s wording, length, structure and intelligibility, as well as the quality of the item’s formulation, the scaling and the items’ and questions’ sequence (Black, 1999). The analysis of psychometric properties of the pilot study’s data concerning the instrument is essential (Fabbriar, Wegener, MacCallum, & Strahan, 1999). Only then is the instrument ready to be administered.

**Method**

**Participants**

The questionnaire was applied in a representative sample of 92 students from three schools of Porto metropolitan area. Sample representativeness was assured by random selection of the participant schools. This sample comprised 46 students from two elementary 9th grade classes and 46 high-school students attending the 12th grade. 25 of these 12th graders were attending Biology and the other 21 were engaged in curricular formats that do not include Biology, namely economics, informatics, humanistic, and arts/design areas. Students from the 9th and 12th grades were chosen to participate in this study for two reasons. Firstly, these are the curricular years where more emphasis is given to biotechnology-related issues, according to the orientations of the Portuguese Ministry of Education (DEB, 2001; DGIDC, 2004). Secondly, since these are the concluding years of elementary and high-school, many of these students end their academic training without any other formal contact with biotechnology.

**Measurement instrument**

The questionnaire’s pilot version was conceived considering instruments previously described (Miles, Ueland, & Frewe, 2003; Cabo Hernandez et al., 2006; Dawson & Soames, 2006; Gaskell et al., 2006; Dawson, 2007; Firmino, 2007; Prokop et al., 2007) and adapting relevant features to the Portuguese educational context. In addition, to better characterize the sample used and to increase the internal consistency of the instrument, new items were formulated. The questionnaire content was decided and assessed upon curricula and textbook analysis, thus assuring its validity. The pilot version consisted in a set of closed and semi-open questions, aiming to provide insights about the knowledge, attitudes, students’ interest towards biotechnology and key sources of information used and trusted. Different scales were developed to assess each of these dimensions. The questions’ wording, especially when translation was required, was carefully considered and the written language was adjusted to the characteristics of the respondents (Oppenheim, 1992). Negative phrasing was avoided and the items were formulated as objectively as possible to minimize bias (Black, 1999).

The questionnaire proposed consists in 14 questions, originally with a total of 65 items organized into 6 groups: knowledge (3 questions, 17 items), attitudes (2 questions, 18 items), interest (2 questions, 4 items), comprehension of news about biotechnology (1 question, 1 item), sources of information about biotechnology used and trusted (2 questions, 15 items) and risk and benefit perception (3 questions, 10 items). Despite its association with knowledge, the comprehension of news about biotechnology depends on how the information is divulged and therefore this item was isolated in its own
category. The risk and benefit perception group comprises questions addressing behavioural intent regarding ethical and controversial issues, to assess the students' decision making capacity. Five factual data questions and one question concerning whether the students found the questionnaire interesting were added to better characterize the population and to enquire the receptivity towards this kind of approach, respectively. Five point Likert type scales were developed for each question, except for the questions in the knowledge section, that consists in a multiple choice question, a list of options and a True/False/Don't know question, and one question aimed at determining sources of information used, that presents a list of options to choose from. A don't know option was included in the True or False question to minimize social desirability bias (Black, 1999; Brace, 2008), and to assess the truthfulness of the students' answers.

Procedure

Data collection

The questionnaires were administered over a three month period, from October to December 2008, during class periods, without imposing time constraints, under the supervision of a teacher and/or the investigator. The time students took to complete the questionnaire was registered. Students were instructed to ask any questions resulting from difficulties in interpreting the questionnaire and regarding any words or concepts that might be unclear to them.

Data analysis

The data collected throughout the pilot study was codified according to a previously defined guideline, recorded and cleansed using the Statistical Package for the Social Sciences (SPSS) software version 17.0. After descriptive and missing values analysis of all the items in the questionnaire, the ones assessed by Likert type scales were subjected to an exploratory factor analysis (principal component analysis with varimax rotation) within their given dimension. Subsequently, by determining the Cronbach's alpha value for each factor identified, a reliability analysis was performed. The number of factors to retain following factor analysis was decided according to the Kaiser criterion (eigenvalues greater than 1) and the scree test (Hayton, Allen, & Scarpello 2004; Costello & Osborne, 2005). Items loading below 0.40, displaying low communality (below 0.40), cross-loading, freestanding or decreasing the scale's internal consistency were excluded from the analysis (Sharma, 1996; Fabrigar et al., 1999; Costello & Osborne, 2005; Hogarty, Hynes, Kronrey, Ferron, & Mumford., 2005). By comprehensively introducing some modifications in the pilot version, the final version of the questionnaire was obtained. It is worth to mention that although the items in the knowledge section and in the question regarding sources of information used were not subjected to factor analysis, both categories are included in the final version of the questionnaire. Item retention for these categories was based on the analysis of missing and ambiguous answers and on the results obtained for the scales subjected to factor analysis.

Results and Discussion

During the pilot study students did not identify major constraints that could justify any modifications regarding the content and structure of the questionnaire. Although there were three complaints about its length, all the students took less than 20 minutes to complete the whole set of questions. The maximum of missing values per item registered for the pilot sample (n=92) was 4.3%, so all the items were considered for subsequent analysis. The missing values were imputed using the series mean method in SPSS, given they were very limited and random (Huisman, 2000; Batista & Monard, 2003; Paul, Mason, McCaffrey, & Fox, 2008).

Factor analysis results, such as the ones summarized in Table 1, lead to the introduction of some changes in the questionnaire's structure, and justified the need to be aware of aspects that influence the interpretation of the data gathered. The total number of items in the questionnaire was reduced by eliminating three items that appeared to be redundant: (i) Genetically modified organisms can endanger the environment (answer True/False/Don't Know) (ii) Rate your agreement towards the following sentence - the ingestion of genetically modified foods has adverse effects on humans (1-3 totally disagree to
5-I totally agree), and (iii) Rate your interest in participating in information campaigns about genetically modified organisms (1-I am not interested at all to 5-I am very interested). It also became evident that a rearrangement of the items was necessary. The attitudes and the risk and benefit perception sections of the pilot version of the questionnaire were combined and their items re-structured according to the tri-partite model of attitudes (Klop & Severiens, 2007), in a section intended to assess the cognitive, affective and behavioural components of attitudes. Additionally, two items, one from each of these two initial sections, were combined into a new category, which factor analysis showed to be consistent with only one factor: importance of biotechnology to the quality of life. The sample’s adequacy to factor analysis was confirmed by the Kaiser-Meyer-Olkin test results, except for the scales addressing the affective component of students’ attitudes and the importance they give to biotechnology (KMO≤0.50). For these scales there seems to be a disperse pattern of correlations among variables, suggesting that a larger sample may yield better results (Sharma, 1996). Nevertheless, for each scale there is a significant ($p<0.05$) correlation between the variables tested, as it was demonstrated by the Bartlett’s Test of Sphericity. Hence, it was decided to keep the factor structures identified for these scales throughout the main study to assess the effect of the sample’s size on its adequacy. Hence, it was decided to keep the factor structures identified for these scales throughout the main study to assess the effect of the increase in the sample’s size on its adequacy. The best factor structures identified by factor analysis for attitude’s cognitive component and trust in sources of information, exclude many of the items initially proposed for those scales (Table 2). However, failing to analyze such items would result in an important loss of information. Therefore, to improve the characterization of the student population, it was decided to include those items along with the factors identified for each scale in the final version of the questionnaire (Table 2).

From the reliability analysis carried out (Table 1), not all the Cronbach’s alpha values obtained are sufficiently robust, scoring below 0.60 (Hair, Anderson, Tatham, & Black, 1998; Wasserman & Bracket, 2003), which may be due to the low number of items for the factors considered (Costello & Osborne, 2005). Although the number of items could be increased, the characteristics of the target population and the feedback obtained during the pilot study rendered a longer questionnaire inappropriate (Oppenheim, 1992; Dörnyei, 2002). At this point, it may be expected that the tendencies observed would be strengthened by increasing the sample size during the final study.

Exploratory factor analysis is an approach that requires many decisions and can produce misleading results (Fabrigar et al., 1999). Therefore, it is necessary to acknowledge that there are many factors affecting its outcomes, such as the design of the study, its aims and the data’s properties (Costello & Osborne, 2005). In this study, our goal was to produce an adjusted instrument that can allow to reliably characterize the studied student population. Despite acknowledging that additional iterations would be beneficial to improve the consistency of the questionnaire’s factor structures, it was decided to use the instrument in its current form during the main study. The validation process described proved to be of the utmost importance to emphasize aspects that were not obvious in the study’s design.
### Table 1. Scales’ factor structure based on exploratory factor analysis (principal component analysis with varimax rotation) and reliability analysis. Coefficients below 0.30 were suppressed.

<table>
<thead>
<tr>
<th>Scale</th>
<th>Factor Analysis</th>
<th>Reliability Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>KMO</td>
<td>Identifiable factor</td>
</tr>
<tr>
<td>Classical applications</td>
<td>0.567</td>
<td>1,186 14,882</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Agro-food applications</td>
<td></td>
<td>1,507 18,842</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bio-medical applications</td>
<td>2,518 31,472</td>
<td>- Utilization of genetically modified cows for the production of medicines for human use</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Embryonic cell utilization</td>
<td>0.449</td>
<td>1,313 32,836</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control capacity</td>
<td>1,171 29,267</td>
<td>- The labeling of transgenic food should specify whether the food or any of its ingredients is genetically modified</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Buying intent</td>
<td>0.624</td>
<td>2,377 36,618</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Access to genetic information</td>
<td>1,118 18,637</td>
<td>- Do a genetic test for medical diagnostic use</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interest</td>
<td>0.674</td>
<td>2,375 59,373</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Industry</td>
<td>0.661</td>
<td>2,106 42,113</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Importance</td>
<td>0.500</td>
<td>1,395 69,740</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Table 2. Questionnaire structure following validation. The factor structure identified for each scale is highlighted in bold.

<table>
<thead>
<tr>
<th>Section / Scale</th>
<th>Example of questions used</th>
</tr>
</thead>
</table>
| Factual questions School you are attending; Grade you are attending; Age; Sex; Course you are attending. | Knowledge assessment  
Q1: Biotechnology can be defined as a set of processes... (Select the option you most agree with)  
(i) ... in which recombinant DNA technology is used; (ii) ... applied to investigation and product development; (iii) ... which involves cell and tissue culture; (iv) ... by which genetically modified organisms (GMOS) can be developed.  
Q2: From the biotechnological applications listed, select the one(s) you know.  
Production of medicines and vaccines; Production of hormones; Production of organic products, such as milk or yogurt; Recovery of contaminated soils using genetically modified bacteria; Waste treatment; Production of insect and pesticide resistant plants; Utilization of plants with industrial purposes, namely the production of cosmetics, plastics and fuels; Production of amino acids and vitamins.  
Q3: Answer the following questions using True/False/I don’t know.  
The ingestion of genetically modified foods can induce gene alterations; Cloning and genetic engineering are identical processes; It is impossible to transfer genes from plants to animals; Genetically modified organisms contain dangerous chemicals; ... |
| Importance assessment | Q4: How important do you think biotechnology is to the quality of life? (1-Not at all important to 5-Very important)*  
Q6c: Rate your agreement with the following sentence (1-I disagree completely to 5-I agree completely): Do you agree that future generations will benefit from biotechnology’s medical applications?|^2|  
| Attitudes assessment – cognitive component | Q5: Rate your approval towards the following activities (1-1 do not approve it at all to 5-1 approve it completely).  
Use of yeast in the production of bread, wine and beer; Use of yeast in animal food production; Use of genetically modified organisms in waste treatment; Improvement of the growth of plant in saline environments by altering their genes; Treatment of genetic disorders by embryonic gene manipulation; Treatment of genetic disorders by human gene manipulation; Insertion of plant genes into animals; Utilization of genetically modified cows for the production of medicines for humans; Production of pesticide resistant plants by gene manipulation; Genetic modification of tomatoes to make them more slowly and have a longer shelf life; Use of insulin produced by bacteria; Organ transplant from transgenic animals to humans; Medical treatments through human cloning. |
| Attitudes assessment – affective component | Q6: Rate your agreement with the following sentences (1-1 totally disagree to 5-1 totally agree).  
It is our duty to authorize investigation that may lead to the development of more efficient medical treatments, even if it implies using embryonic stem cells; The labels of transgenic food should specify the food or any of its ingredients is genetically modified; It is wrong to use embryonic stem cells in biomedical research, even if it may contribute to the development of medical treatments; Each of us is capable of determining our intake of transgenic foods. |
| Attitudes assessment – behavioural component | Q7: How often would you... (1-1Never to 5-Almost)  
Buy transgenic foods if they were available in supermarkets; Buy medicines obtained by genetically manipulation.  
Q13: How often would you... (1-1Never to 5-Almost)  
Do a genetic test for medical diagnosis; Give the police access to your genetic information; Buy transgenic foods if they were healthier than other foods; Buy transgenic foods if they were less expensive than other foods. |
| Interest assessment | Q8: Rate your interest towards biotechnology (1-I am not interested at all to 5-I am very interested)!  
Q9: How often do you... (1-1Never to 5-Many times)  
...listen to news about biotechnology; ...read articles or watch TV shows about technology; ...search the web for subjects related to biotechnology. |
| Comprehension of news assessment | Q10: Rate your difficulty in understanding news about biotechnology (1-Very low to 5-Very high). |
| Assessment of sources of information used | Q11: From which of the following sources do you most commonly obtain information about biotechnology?  
TV; Radio; Newspapers, Magazines; Scientific magazines; Internet; Textbooks; Teachers; Friends; Family; Others. |
| Assessment of trust in sources of information | Q12: Rate your trust in the following sources of information about biotechnology (1-I do not trust it/them at all to 5-I trust it/them completely):  
Media; Scientific magazines; Pharmaceutical industry; Agro-food industry; Health industry; Governmental agencies; Universities; Scientists; Internet; Environmental organizations; Consumer rights organizations; European Community; Medical doctors; Politicians. |

*Note. Items identified with the same letter (a,b,c,d,e,f,g,h,i,j,k) contribute to the same factor.
Conclusions and Implications

We believe that the quality of the data gathered through an inquiry-based methodology can be improved by an adequate validation of the instruments used. In addition, this procedure may be fundamental to assess and optimize the reliability and validity of those instruments and the results they provide. When developing new questionnaires, analysis of psychometric features can help reduce bias introduced by the author’s own expectations towards the students’ answers. Furthermore, it can unveil intrinsic and unpredicted conditioning factors. Validation also becomes crucial when planning to use already existing instruments, since the population to be assessed presumably differs from the ones for which those instruments were originally intended.

Overall, the work carried out underlines a guideline to similar diagnostic studies. A validation strategy was tested and its outcomes discussed, demonstrating that the investment in these time consuming procedures can improve the quality of the data gathered through quantitative assessment methodologies.

Acknowledgements

The authors are grateful to all the participant students and teachers and to Catarina L. Santos for helpful comments and suggestions on the manuscript. Maria João Fonseca is supported by the FCT fellowship SFRH/BD/37389/2007.

References


