SQP Users' Manual

2015

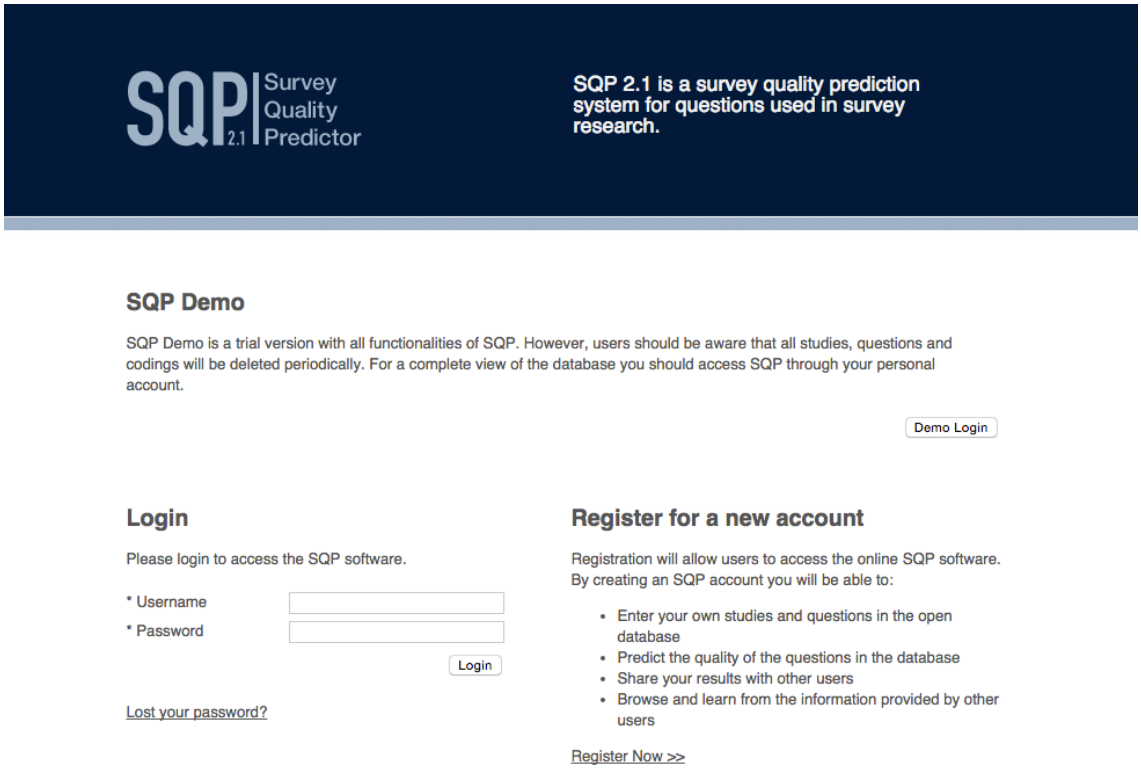
To cite this document: Survey Quality Predictor (2015). SQP Users' Manual. Barcelona, Universitat Pompeu Fabra.

TABLE OF CONTENTS

The Survey Quality Predictor (SQP) software is used to predict the measurement quality of survey questions. The SQP quality prediction is based on an analysis of the relationship between the quality of questions obtained from past Multitrait - Multimethod (MTMM) experiments[1] and their characteristics[2]. SQP is online software that allows researchers to add a question, code its characteristics for more than 20 countries[3], and obtain a quality prediction. For more detailed information regarding the characteristics of the questions, go to Saris and Gallhofer (2014)[4] and to the SQP Coding instructions (2015)[5].

SQP consists of an open database of survey questions, quality predictions and MTMM estimates. SQP users can be participant in the development of this database, as they are able to add survey questions, see other users' questions, create their own quality predictions and also see other users' predictions. Because SQP is open software, it is strongly recommended using the SQP Demo for testing the software and learning the coding procedure of the characteristics. The Demo interface is easily accessible from the login page (see Screen 1).

Screen 1: SQP login page



[1] MTMM experiments are conducted in order to determine the quality of survey questions.

[2] Formal and linguistic characteristics such as: domain, social desirability, balance of the request, response type, presence of neutral category, respondent instructions, number of nouns in the request, use of showcard, presence of an interviewer, etc.

[3] The current version of SQP is able to provide predictions for the following countries in combination with any language: Austria, Belgium, Czech Republic, Denmark, Estonia, Finland, France, Germany, Great Britain, Greece, Ireland, Netherlands, Norway, Poland, Portugal, Slovenia, Slovakia, Spain, Switzerland, Sweden, Ukraine and United States. The effect of the language is captured by the coding of the linguistic characteristics of the questions. However, users should take into account that the effect of the linguistic characteristics are measured by the use of the official languages in those countries.

[4] Saris, W. E. and Gallhofer, I. N. (2014). Design, evaluation and analysis of questionnaires for survey research. Second edition, Hoboken, Wiley.

[5] Survey Quality Predictor (2015). SQP Coding Instructions. Barcelona, Universitat Pompeu Fabra.

1. Using the SQP Demo

The SQP Demo provides users with the complete set of features in SQP. Thus, Demo users are able to create new questions, check existing questions, characteristics' codes and quality predictions, and code the characteristics of their own or existing questions to obtain their own quality predictions. The only two differences between the SQP Demo and SQP are:

- In the SQP Demo all created studies[6], questions and characteristics' codes will be deleted periodically.
- The complete database of questions cannot be found in the SQP Demo. This is only available for registered SQP users.

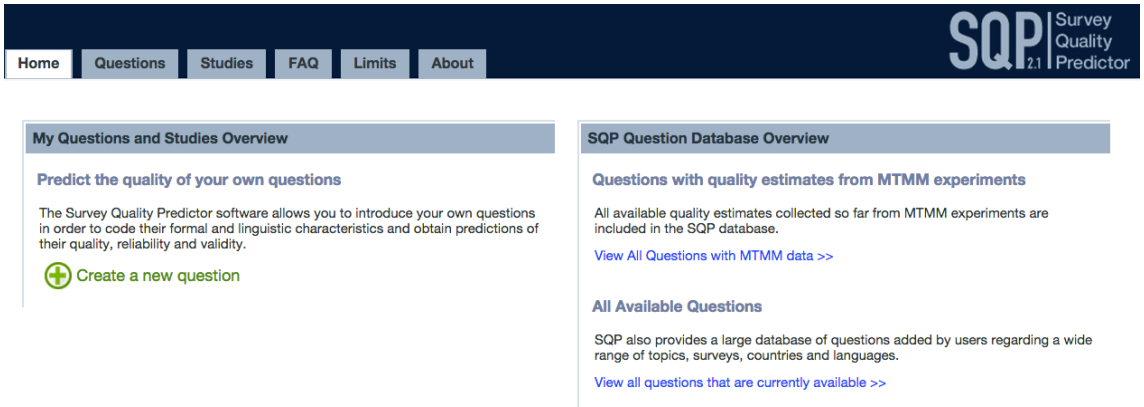
However, the SQP Demo does provide a study which contains a sample of the type of survey questions that can be found in the SQP database. In this study, named as 'SQP Illustration', the users can find a set of questions with MTMM estimates and authorized quality predictions[7].

Once the Demo users are familiar with the use of the software and the coding procedure of the characteristics, it is recommended registering with SQP using a personal username and password.

2. Getting started

Once logged in as registered SQP users, users are directed to the SQP Home page. As illustrated in Screen 2, the SQP Home page provides a good overview of the main features of the software.

Screen 2: SQP Home page



Screen 2 shows that in SQP three main options are available. Users can: 1) Create a new question, 2) view the questions that were part of a MTMM experiment, and 3) view all available questions in the database. Below, these three features are described.

[6] A study within SQP, groups a set of survey questions under the generic and meaningful title. It is strongly recommended naming the SQP studies under the name of the research project, the questionnaire or the survey agency or organization (e.g. ESS Round 1 refers to all questions in the first Round of the European Social Survey).
[7] The SQP team can authorize quality predictions.

2.1. Creating a new question in SQP

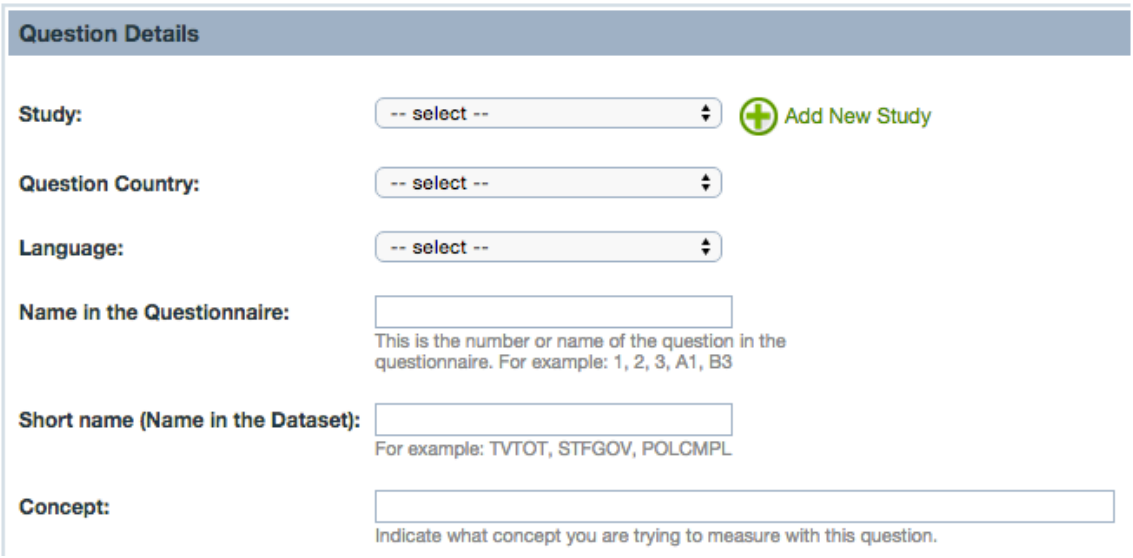
Following the link 'Create a new question' on the SQP Home page (see Screen 2), users are able to specify the details of the new question they that want to add. These details, as shown in Screen 2.1.1, refer to the study to which it belongs, the language in which it is written and the country in which the survey question is asked. The countries for which SQP is able to provide a prediction are limited. The list of countries for which the current version of SQP is able to provide a prediction is the following:

Austria, Belgium, Czech Republic, Denmark, Estonia, Finland, France, Germany, Great Britain, Greece, Ireland, Netherlands, Norway, Poland, Portugal, Slovenia, Slovakia, Spain, Switzerland, Sweden, Ukraine and United States.

Therefore, if your question belongs to a country not on the list, SQP will require you to specify a 'Prediction country' for which you will be able to obtain a quality prediction. It is suggested you select as a second country the one from the list which is most similar in terms of linguistic and cultural characteristics. With regard to the selection of a study, it is recommended that users choose a study which is meaningful for the question created. Preferably, users can choose a study which has already been created by him or her, rather than adding questions to studies, which are neither his or her own or related to their question. If the question the users want to create belongs to an already existing study, users would be able to find it in the database and code it. For testing non-meaningful or serious questions, it is suggested that the SQP Demo is used, which was created for this purpose.

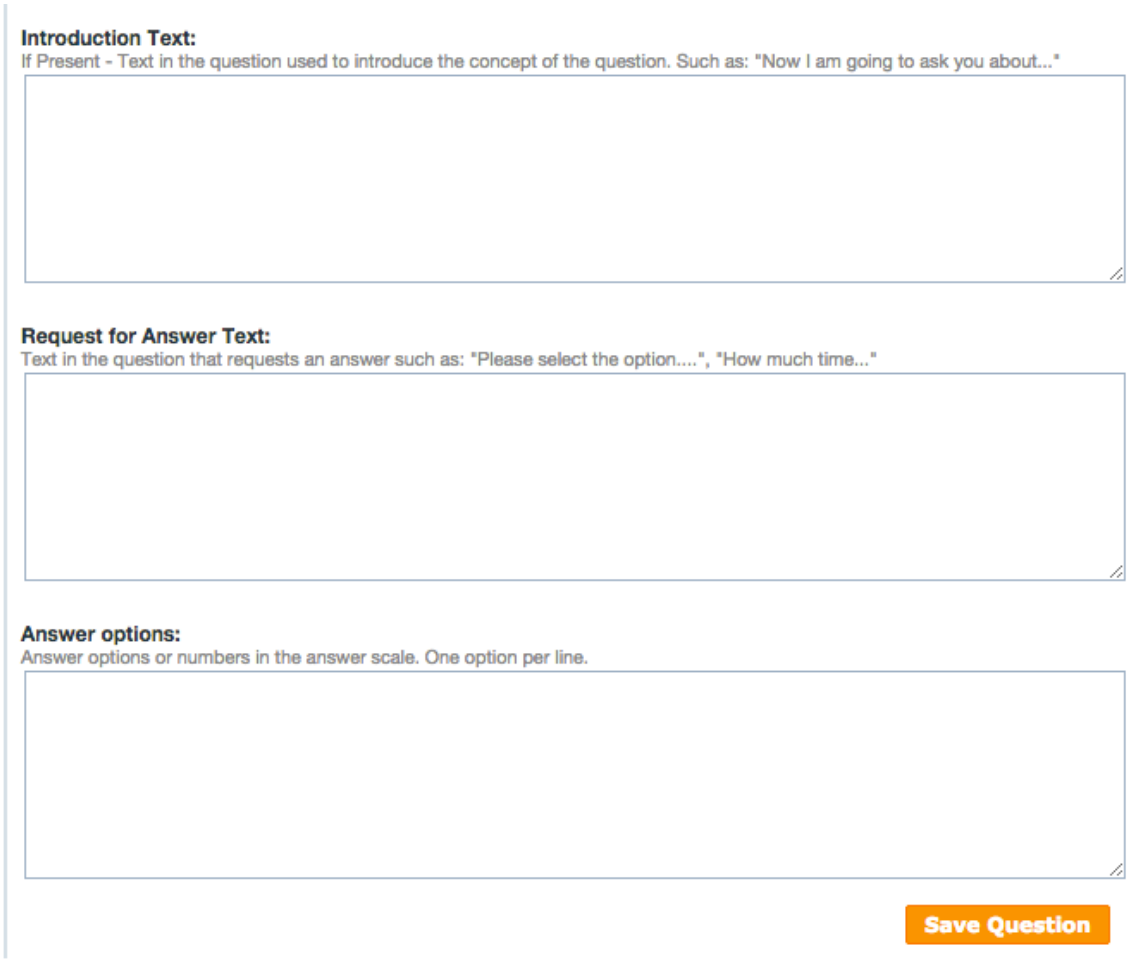
Furthermore, users are able to specify a name for the question, which can be the name of the question in the questionnaire, a short name, which can be the name in the dataset, and the concept that is measured.

Screen 2.1.1: New question details



To add a new survey question, users have to specify the text of the question in the correct boxes. SQP provides the users with three different boxes for the question text: 1) the 'Introduction Text' box, 2) the 'Request for Answer Text' box and 3) the 'Answer Options' box.

Screen 2.1.2: Introduction of a new question text



As illustrated in Screen 2.1.2, in the first box users have to write the introduction text of the question, if there is any. Introductions usually serve to introduce the concept of the next question(s). For example:

"Next I will ask you some questions about politics."

In the second box, users have to write the text of the question, or as SQP calls it, the request for an answer[8]. This text should only include the text that is actually read to or by the respondent, thus, interviewer instructions should be left out. For example, a request for an answer would be:

"Could you tell me to what extent you agree or disagree with the way democracy works? Use this card to answer."

Finally, in the third box, the answer options have to be introduced. If it is a close-ended response scale, users have to introduce one option per line. However, if it is an open-ended question, users should leave the box empty. Continuing with the example, the answer options could be introduced as follows:

[8] SQP refers to questions to the text that involves an introduction, a request for an answer and the answer options. Thus, an open-ended question without an introduction will be equal to its request for an answer.

"1. Completely disagree

2. Disagree

3. Neither agree nor disagree

4. Agree

5. Completely agree"

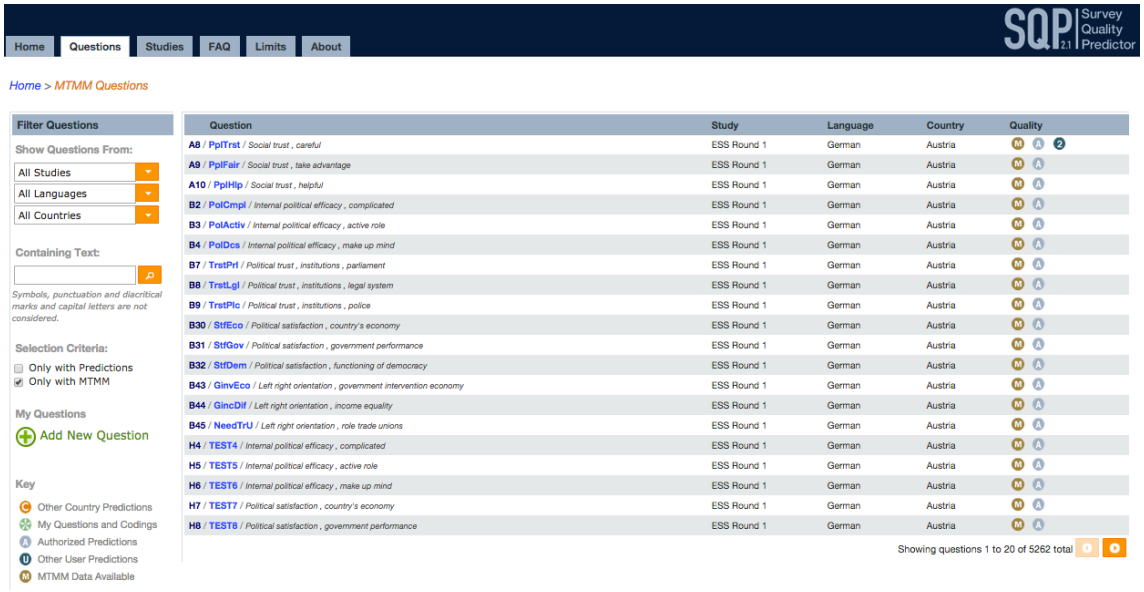
Only if the 'Don't Know' option is explicitly provided, users should add the 'Don't Know' as an extra option. However, if the 'Don't Know' option is implicitly used or not used at all, then it should not appear in the 'Answer Options' box. In any case, it will be later coded as a formal characteristic of the response scale.

2.2. Questions with MTMM estimates

The MTMM estimates available in SQP refer to the estimated results from MTMM experiments carried out in different surveys[9] over the last decades. These results are not obtained through the SQP prediction system. Instead, most of them are actually used in SQP as part of the analysis together with the characteristics of the questions, which allows the quality of new questions to be predicted.

If users are interested in checking the questions available with MTMM estimates, they can access them through the SQP Home page by clicking 'View all questions with MTMM estimates' (see Screen 2). This link will provide users with the list of survey questions with MTMM estimates, as in Screen 2.2.

Screen 2.2: List of questions with MTMM estimates in the SQP database



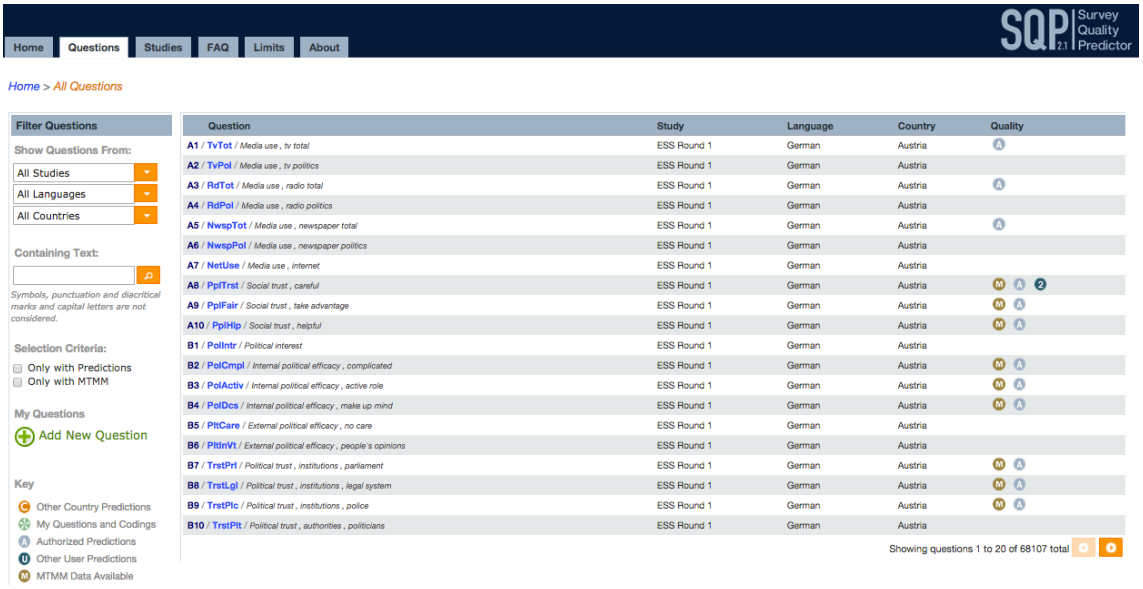
All MTMM questions are identified by the key (M) and they also have authorized quality predictions. Authorized predictions can be identified by the key (A).

[9] MTMM experiments carried out in surveys such as: European Social Survey (ESS), Eurobarometer, Dutch Gallup Organization, Study of Michigan Generations Panel survey, etc.

2.3. All available questions in SQP

The list of all available questions in the database is shown on the SQP Home page, as presented on Screen 2.3, by clicking on the link 'View all questions which are currently available' (see Screen 2).

Screen 2.3: List of all available questions in the SQP database



The list of questions shows the name of the questions (in navy blue) and the short name (in a lighter blue color). It also provides information regarding the concept measured, the study it belongs to, the language, the country, and the type of quality information available. In the 'Quality' column, users can see if the question belongs to an MTMM experiment (M), if the quality prediction belongs to the users themselves ( ), or to another user (U), and/or if it has been authorized (A) by the SQP team.

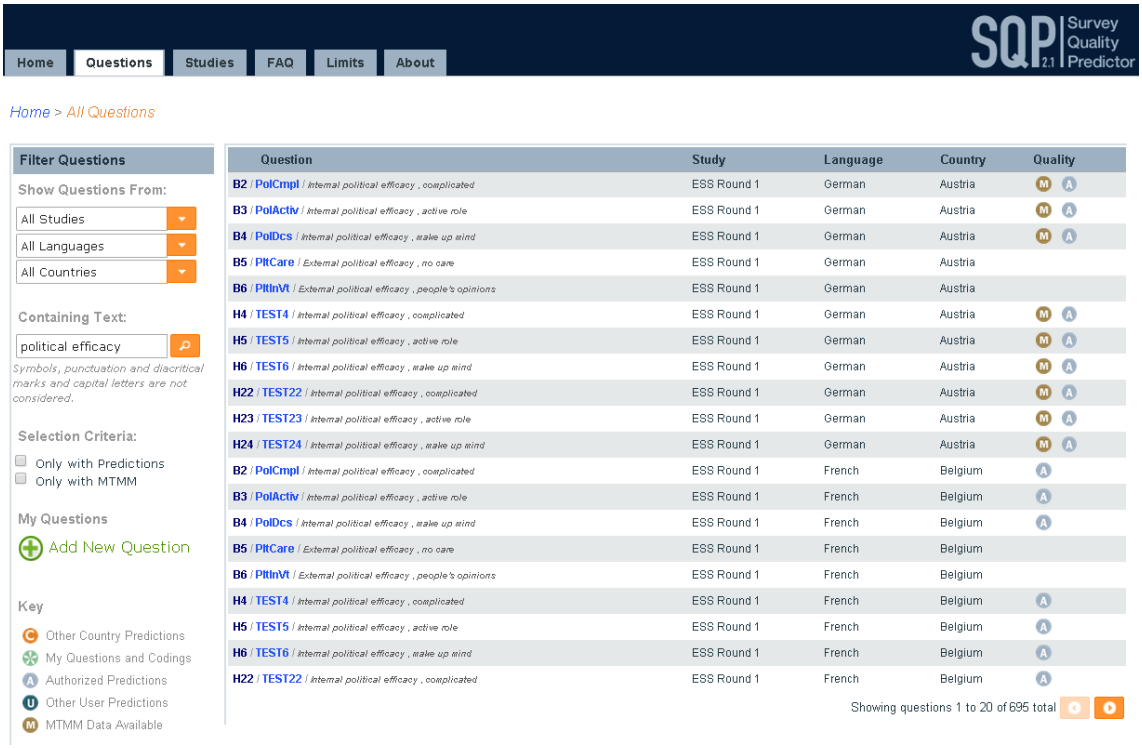
An authorized prediction implies the coding of the formal and linguistic characteristics being done by trained native speakers under the control and supervision of the SQP team.

Furthermore, the 'Quality' column indicates if the question prediction belongs to a different country than the one presented for the question, i.e. a 'Prediction country'. This is indicated by the key C.

3. How to find your question of interest in SQP

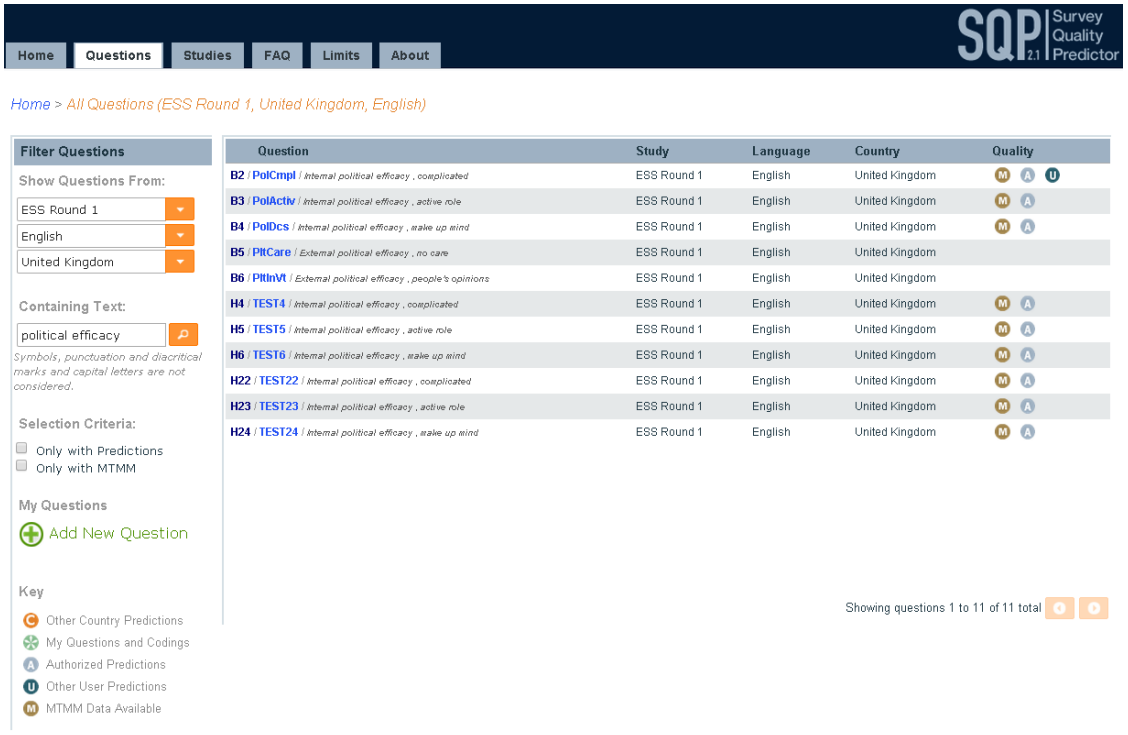
Once on the list of questions screen, if users are interested in a specific topic, they should use the 'Containing text' filter option. The 'Filter Questions' options are placed in the column on the left hand side of the question list (see Screen 2.3). The 'Containing text' box looks for the wording in any part of the question (i.e. the name, the short name, the concept or the question texts). This leads users to a shorter list of questions containing the selected text for the languages, countries and studies for which it is available. Screen 3.1 illustrates that by searching for the word "political efficacy", where the complete list of more than 65,000 questions (see Screen 2.3) is reduced to 695 questions.

Screen 3.1: Containing text filter



However, if users already know of the existence of a specific question in a study or are only interested in a certain language or country, then they can add to the filter options the study, the language and/or the country to obtain an even shorter list. For instance, imagine that users are interested in a question regarding “political efficacy” asked in the United Kingdom and users know that such a topic was asked in ESS Round 1. Users should select from the options: ‘All Studies = ESS Round 1’, ‘All languages = English’ and ‘All countries = United Kingdom’, as in Screen 3.2. With these details, the list has now been reduced to 11 questions.

Screen 3.2: Using filter options in SQP

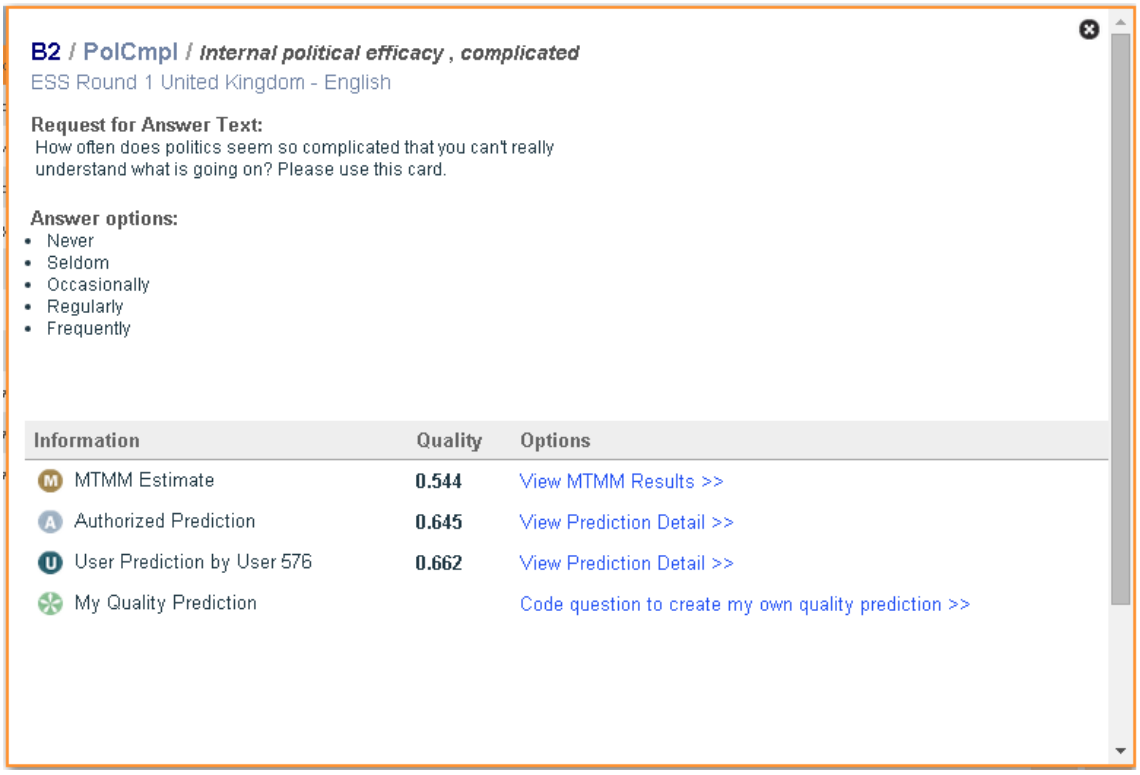


This list shows that 9 out of the 11 questions belong to MTMM experiments. This means that information about the results of these experiments is available for each of these questions. These 9 questions have also already been coded and authorized by the SQP team, meaning that a trustworthy prediction is available for the users. Furthermore, the first question in the list, question B2, has a quality prediction made by another user.

3.1. When the question of interest is already coded

If users choose a question that has already been coded, the details of the quality predictions will already be available. For example, by selecting question B2, (in Screen 3.2) users will get the pop-up Screen 3.1.1, in which the question text and a summary of the quality information are presented.

Screen 3.1.1: Predictions available for the users



Screen 3.1.1 provides two different quality predictions and an MTMM estimate. Users will notice that the qualities vary considerably. Firstly, it is important to mention that the MTMM estimates and the SQP predictions do not necessarily have to provide the same quality. Taking into account that the quality estimates obtained through MTMM experiments are the result of a one point in time analysis, and given that the SQP predictions are based on the results of many of these analyses, in general, the SQP quality prediction is a more reliable source than the MTMM estimates. There are only two cases in which it would be advisable to rely on the MTMM estimates rather than the SQP predictions:

1) Agree – Disagree (AD) scales cannot be identified yet in the SQP characteristics (see Limit 4). Thus, although it has been demonstrated that the quality of item-specific (IS) scales is much higher than AD scales (Saris et al., 2010)[10], SQP would not be able to distinguish both.

[10] Saris, W. E., Revilla, M., Krosnick, J.A. and Shaeffer, E.M. (2010). Comparing Questions with Agree/Disagree Response Options to Questions with Construct-Specific Response Options. *Survey Research Methods, 4(1): 61-79. Available at: https://ojs.ub.uni-konstanz.de/srm/article/view/2682*

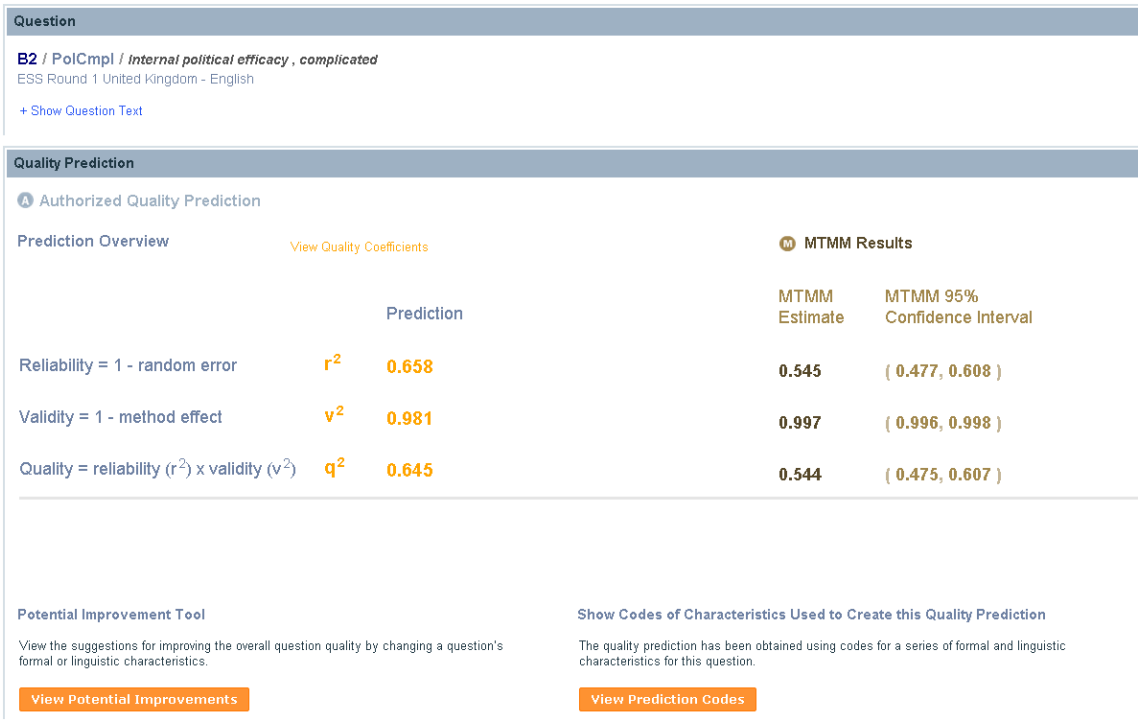
2) If a question in a country deviates considerably from the general trends of the question in the database of all countries. The SQP predictions are based on the general trends of the relationships between the different characteristics of the MTMM questions and the MTMM quality estimates of these questions in all countries.

Furthermore, it is important to remember the distinction between user predictions (U) and authorized predictions (A). The coding procedure of the questions' characteristics is not so simple. Because questions with authorized predictions have been coded by trained native speakers under the control and supervision of the SQP team, the codes and predictions of these questions are the most trustworthy ones.

Whenever the question of interest is coded but not authorized, it is recommended double-checking the coding before using its results, as the coding decisions affect the prediction results. Therefore, users should do the coding carefully. As far as possible, it is recommended doing the coding with two independent coders in order to prevent unintended errors.

Furthermore, users can obtain the estimates and predictions for the reliability, validity and quality by clicking on the link 'View prediction detail' (in Screen 3.1.1). The prediction details for the authorized prediction and MTMM results are presented in Screen 3.1.2.

Screen 3.1.2: SQP quality prediction details and MTMM results

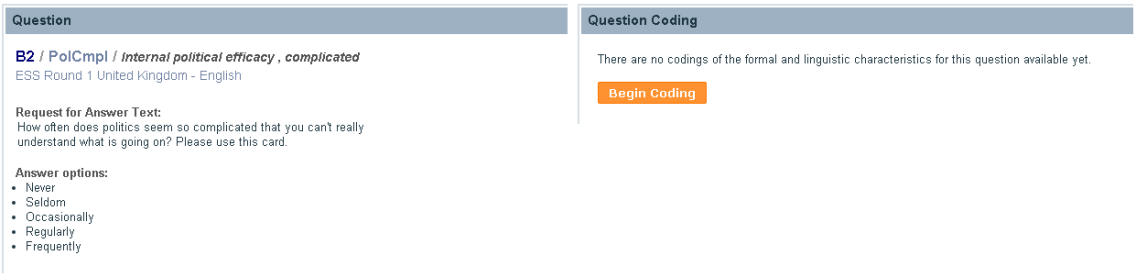


4. Coding the characteristics of a question

In order to start coding the characteristics, users have to select the question of interest and click on: 'Code question to create my own quality prediction' (the last link in Screen 3.1.1). Even if another user has already coded the question, users will still be able to code it.

The software will show, as in Screen 4.1, the selected question to be coded, including the introduction (if present), the request for answer and the answer options. However, this text is just a reminder and should not be the object of coding. Coders should base their codes on the country questionnaires and showcards (if used). For example, the interviewer instructions and the "Don't Know" option would hardly ever appear on the screens, therefore users should rely on the information in the questionnaires to see whether there are instructions or not.

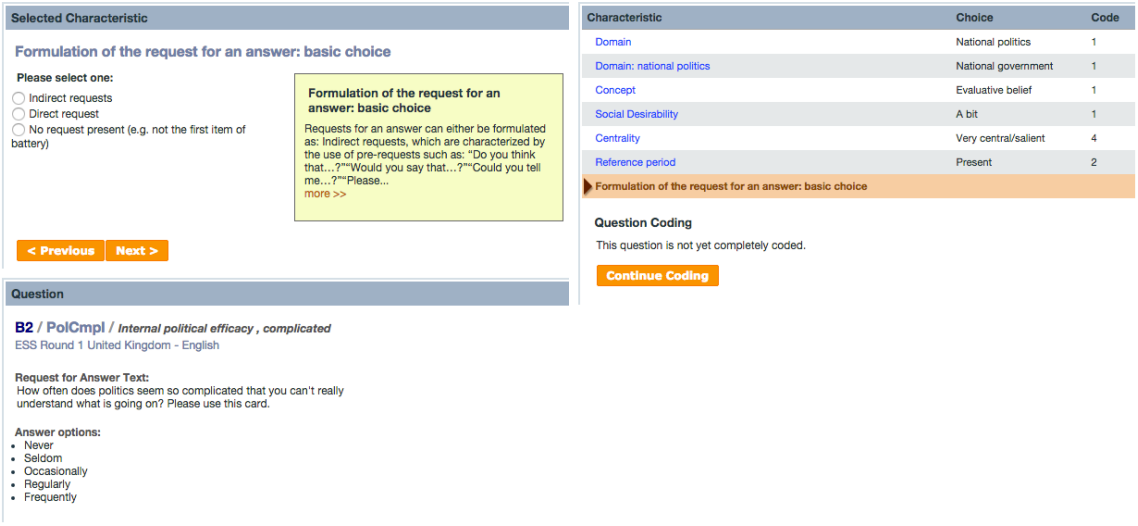
Screen 4.1: Starting the coding procedure



Once the users start the coding process, SQP will provide the set of characteristics that need to be coded based on the question selected and previous coding decisions. For example, if users specify that there is no introduction present, SQP will not ask about the linguistic characteristics of the introduction.

The choices of characteristics appear above the question text, on the left hand side of the screen. To keep track of the users' coding decisions, SQP provides the progress of the coding on the right hand side of the screen. Screen 4.2 is an illustration of what the coding looks like when it is still in progress. In the list on the right hand side, it is indicated that 6 characteristics have already been coded, among which are: the 'Domain' and the 'Concept' of the question. Now users have to indicate in the options above the question text if the request is formulated in a direct or indirect way, or if no request is present in this question.

Screen 4.2: Coding progress



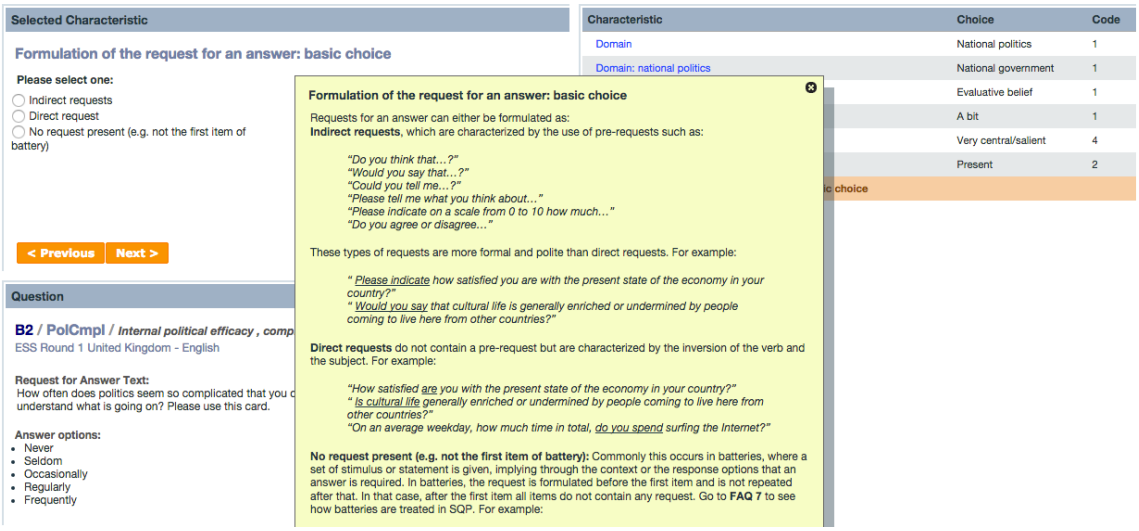
To obtain an overview of the type of characteristics users will be asked to code in SQP, Table 1 provides a short description of the basic characteristics they will encounter while coding. Note that not all characteristics have been introduced in the table because some characteristics depend on the previous coding, e.g. specific characteristics for the type of response options, specific linguistic characteristics if the introduction is present, specific characteristics if a showcard is used, etc.

Table 1: Summary of SQP basic characteristics	
Characteristics	Short Description
Domain	The domain is the topic of the assertion that one wants to measure using this question. It is determined by the research goal.
Concept	The concept that they want to measure should be classified in one of the basic concepts distinguished on the list.
Social desirability	Social desirability is connected with the choice of the domain. Identifies sensitive/delicate/irritable questions, which can bias the responses obtained.
Centrality	Centrality is also directly connected with the choice of the domain. It measures the familiarity of the respondents with the topic.
Reference period	Is again also connected with the research topic. The time period mentioned in the request can be: present, past or future.
Formulation of the request for an answer: basic choice	Identifies if a request is formulated as a direct or indirect request or if there is no request present, which means that the request (in this case the stimuli or statement) belongs to a battery of questions (except for the first item of the battery which will be either a direct or an indirect request introducing the battery).
WH word used in the request	Identifies questions that use words like: 'who', 'which', 'what', 'when', 'where' and also 'how', 'to what extent', 'to what/which degree' or 'whether'. Or translations of them in other languages.
Request for an answer type	Identifies if a question is formulated in an interrogative, imperative or declarative form.
Use of gradation	Identifies requests that indicate responses that can be ordered from low to high or from high to low.
Balance of the request	Identifies leading questions. A request is Balanced when it contains both possible answer poles and Unbalanced when just one pole is mentioned.
Presence of encouragement to answer	Identifies leading questions. A request is Balanced when it contains both possible answer poles and Unbalanced when just one pole is mentioned.
Emphasis on subjective opinion	Identifies an emphasis on the opinion of the respondent about something, like: 'Please give us your opinion about…', 'According to you…', 'What do you think about…', etc.
Information regarding the opinion of other people	Identifies when opinions of other people are given in the request, like: 'Some people are against nuclear energy while others are in favour of it…'
Use of stimulus or statement in the request	Identifies batteries of questions. A stimulus in a question can be a noun or a combination of nouns. A statement in a question consists of complete sentences.
Absolute or comparative judgement	Identifies if the respondent has to compare two events or things. For example: 'Are you feeling better than last year or not?'
Response scale: basic choice	Identifies what types of answer options are provided: 1. Categories, 2. Frequencies, 3. Yes/ No scales, etc.
Response scale characteristics	Number of categories, frequencies and maximum possible value
	Labels of the categories
	Labels with short or long text
	Order of the labels
	Correspondence between the labels and the numbers of the scale
	Theoretical range of scale bipolar/ unipolar

	Range of the used scale bipolar/ unipolar
	Symmetry of the response scale
	Neutral category
	Number of fixed reference points
Don't know option	Identifies whether there is a Don't know option.
Interviewer instruction	If an interviewer administrates the survey, interviewer instructions will often be present, regarding which card to use or how to continue.
Respondent instruction	An instruction to the respondent is often present in imperative requests. However, these instructions explicitly ask the respondent to do something.
Extra motivation, information or definition available	Identifies if there is an extra sentence introducing a motivation, other information or a definition of something.
Introduction available	Identifies the presence of an introduction, which mainly serves to initiate the topic of the request to the respondent.
Linguistic characteristics of the introduction, the request for an answer and the answer scale	Number of sentences in the introduction
	Number of words in the introduction
	Number of subordinate clauses in the introduction
	Request present in the introduction
	Number of sentences in the request
	Number of words in the request
	Total number of nouns in the request for an answer
	Total number of abstract nouns in the request for an answer
	Total number of syllables in request
	Number of subordinate clauses in request
	Number of syllables in answer scale
	Total number of nouns in answer scale
	Total number of abstract nouns in answer scale
Show card used	Identifies the use of Show cards. These are sometimes used during the interview to show the response options or to assist in explaining the question.
Show card characteristics	Horizontal or vertical scale
	Overlap of text categories
	Numbers or letter before answer categories
	Scale with numbers or numbers in boxes
	Start of the response sentence on the show card
	Question on the show card
	Picture provided
Computer assisted	Identifies the mode of data collection: if the interview is computer-based or not.
Interviewer	Identifies the mode of data collection: if it is a personal interview or a self-administered questionnaire.
Visual or oral presentation	It identifies if the questionnaire is self-administered (visual) or interviewer-administered (oral).
Position	Identifies the position of the question in the questionnaire.

SQP provides Coding Help screens, such as the yellow box presented in Screen 4.3, which explain every characteristic in detail. However, for specific questions regarding the coding, coders can send an email to: sqp@upf.edu.

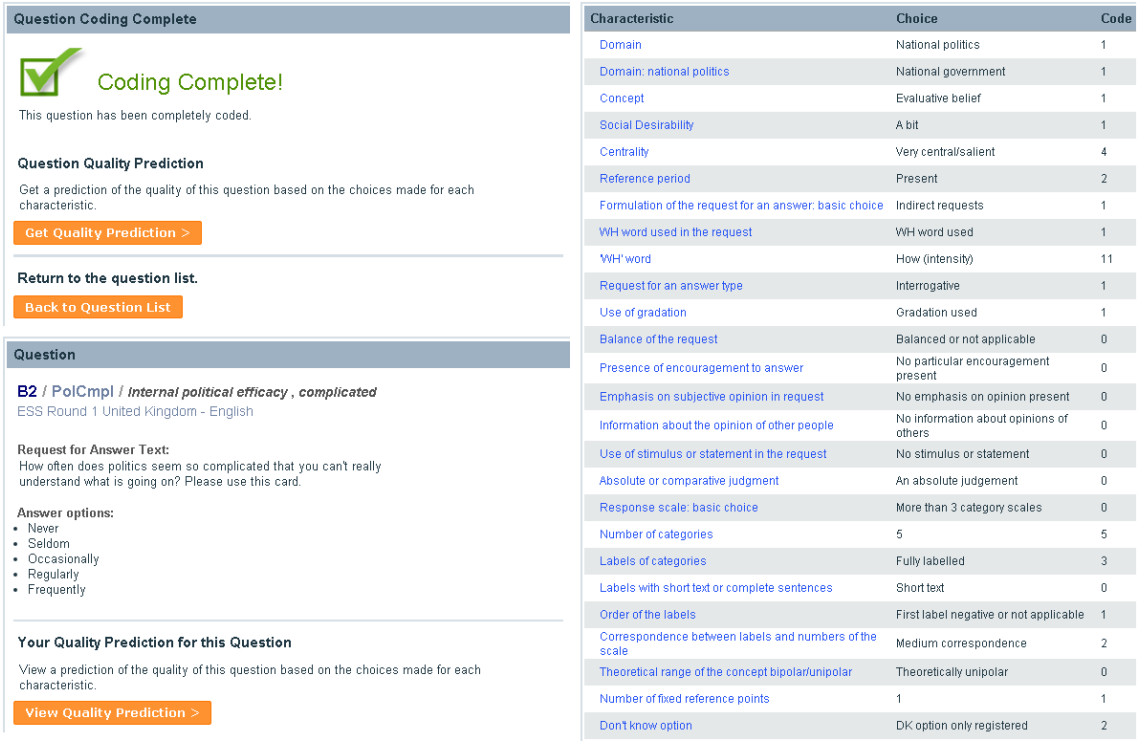
Screen 4.3: Coding Help screens



5. Task finished

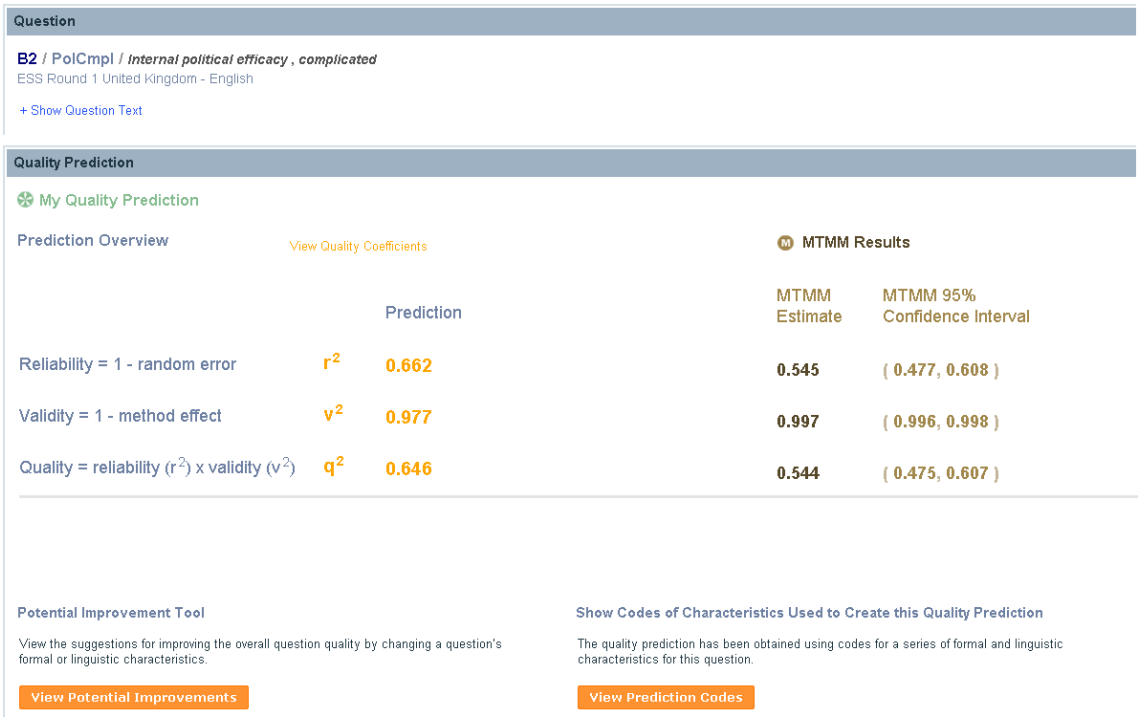
To complete coding the characteristics of a survey question and obtain a quality prediction, users will need to identify between 30 and 60 characteristics of the question from a total of 73, depending on the complexity of the question being evaluated. At first, this may seem rather tedious. However, as soon as users become familiar with the characteristics and their meanings, the coding procedure becomes simpler and faster. As illustrated in Screen 5.1, the software will display a message when the coding is completed.

Screen 5.1: Coding complete



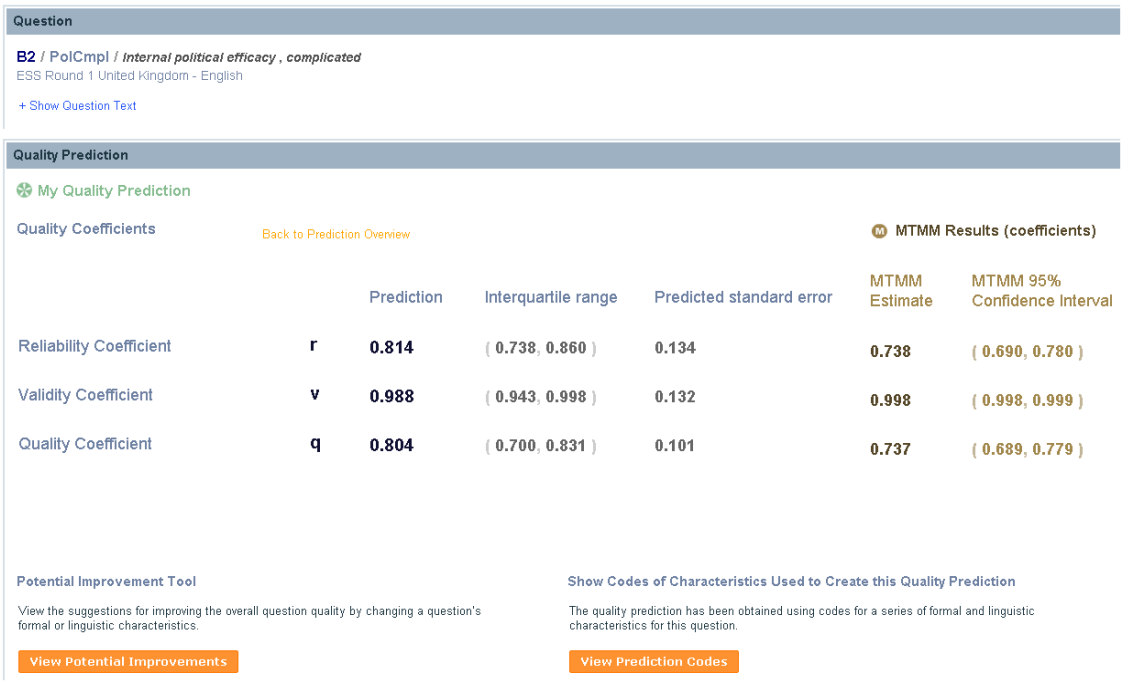
By selecting ‘Get quality prediction’, users can obtain predictions of the reliability, validity and quality of the question (see Screen 5.2). By selecting ‘Back to question list’ users can go back to the screen which displays the list of filtered questions (as shown in Screen 3.2).

Screen 5.2: Prediction overview



In the predictions overview, users will find the information regarding the reliability, validity and quality of the predictions and those from the MTMM results (if available). Moreover, users have the option to ‘View Quality Coefficients’ (see Screen 5.3). These are the square root of the quality indicators on the previous screen.

Screen 5.3: Quality coefficients

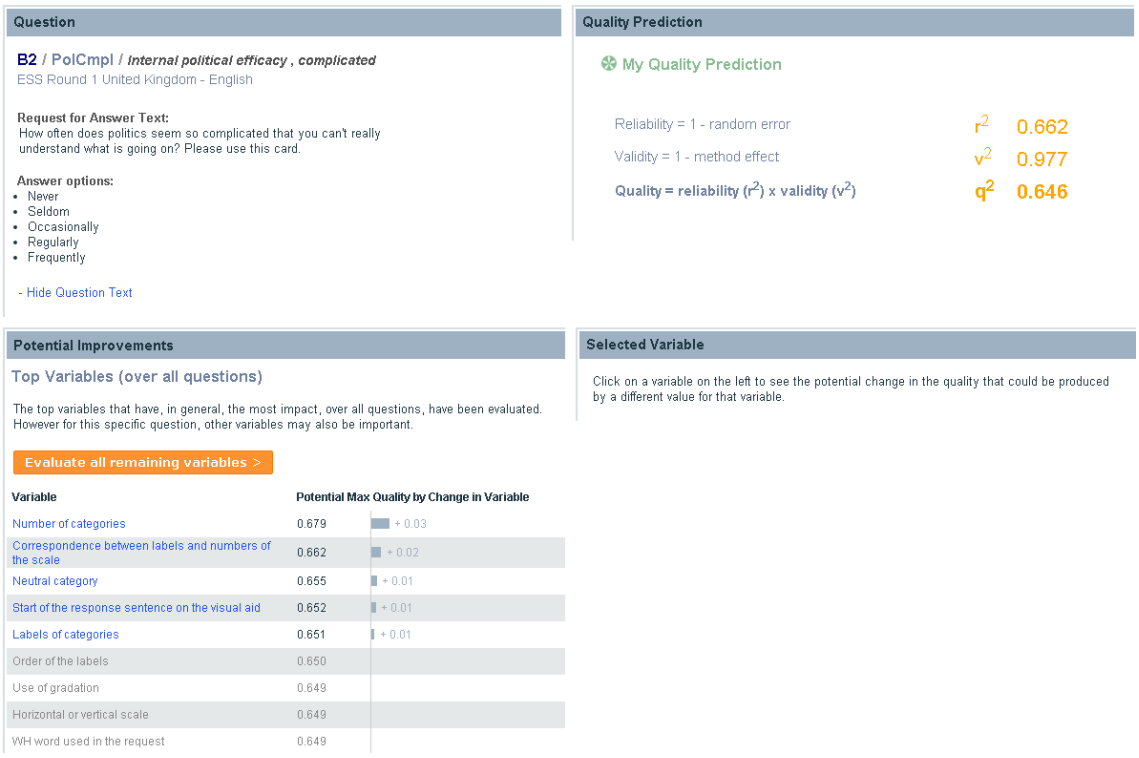


On Screen 5.3, SQP also provides the uncertainty of the estimates by indicating their interquartile ranges and the standard errors.

6. Potential improvements tool

To conclude, SQP not only allows users to obtain the quality prediction of survey questions, but also to view the potential improvements that could be made in the survey question to obtain a higher quality. The 'View potential improvements' tool can be accessed from the link on the bottom left hand side of Screen 5.2 and Screen 5.3. This tool suggests to users which characteristics of the survey question could be corrected in order to obtain a better quality (see Screen 6.1).

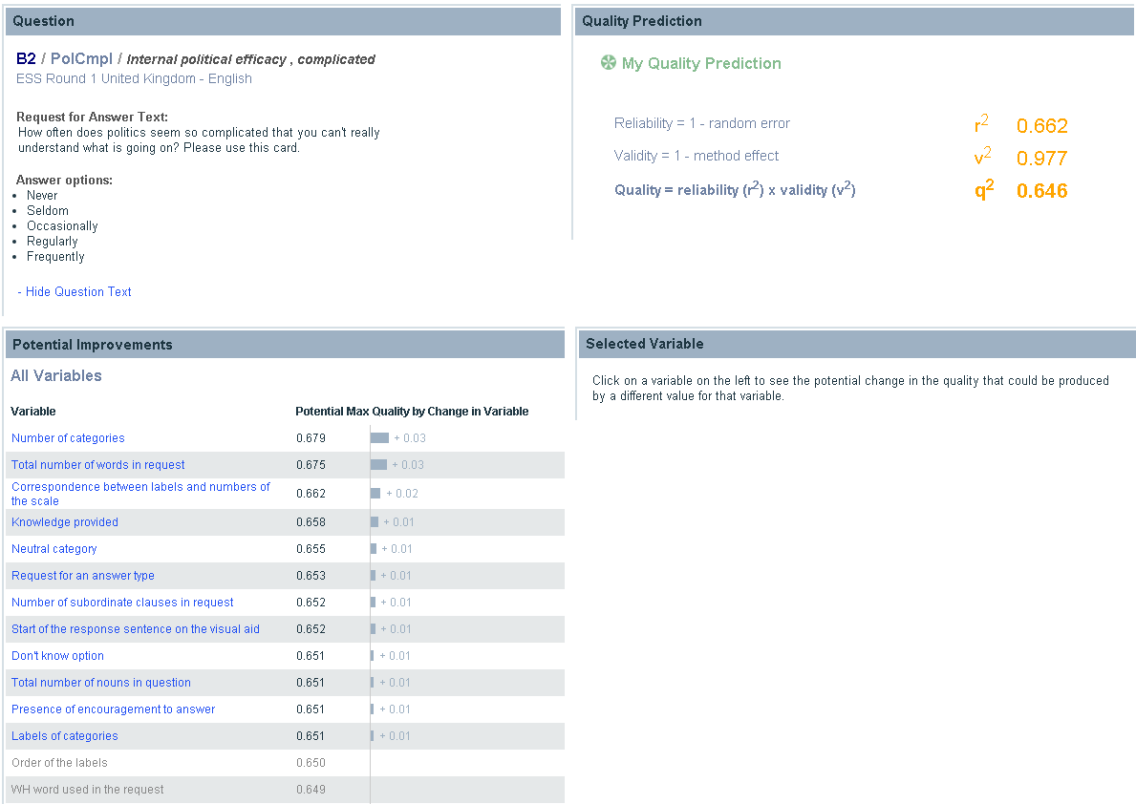
Screen 6.1: View potential improvements tool, 20 first characteristics



In the context of designing a survey question, users should take into account the purpose of the question, and evaluate the tool suggestions critically when deciding whether or not they apply to the survey question.

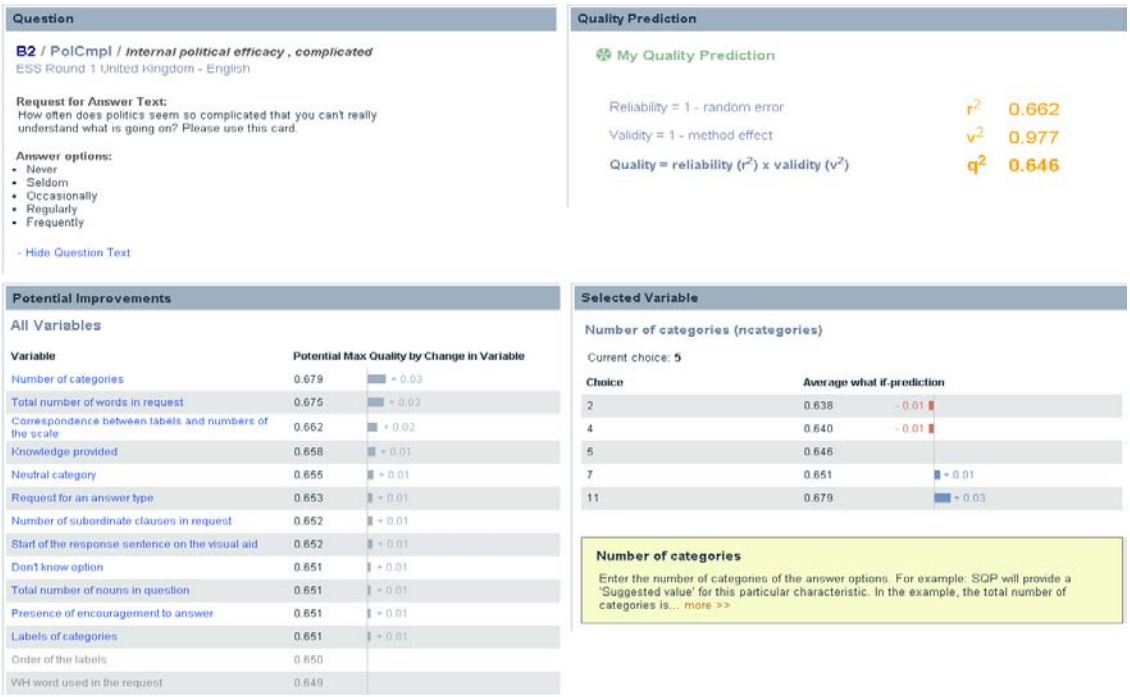
This tool firstly presents a selection of characteristics which are most important for all questions. To obtain the complete list of suggested improvements, click on "Evaluate all remaining variables" (in Screen 6.1). The complete list consists of 63 characteristics that can be used to improve the question (see Screen 6.2).

Screen 6.2: View potential improvements tool, all characteristics



Detailed information about the specific change to be made is obtained by clicking on the characteristic of interest (see Screen 6.3). Then, on the right hand side, users obtain the information for which aspect of this characteristic the largest improvement in quality of the question is available.

Screen 6.3: Potential improvement suggestion



The effects on the quality that SQP calculates are independent of other related changes that the tool suggests. This means that the effects presented do not take into account related characteristics that may also change due to the change being suggested. For instance, increasing or reducing the number of points in a scale may lead to other changes in the characteristics related to the answer options, such as the labels or the presence of a neutral category. These consequences are not immediately taken into account. Therefore, it makes sense to reformulate the question and test the new question in the same way as indicated above.