# A Dynamic Questionnaire to Further Reduce Questions in Learning Style Assessment

Esperance Mwamikazi[1], Philippe Fournier-Viger[1], Chadia Moghrabi[1], Robert Baudouin[2]

[1]Department of Computer Science, Université de Moncton, Canada
[2]Department of Secondary Education and Human Resources, Université de Moncton, Canada
{esperance.mwamikazi, philippe.fournier-viger, chadia.moghrabi, robert.baudouin}@umoncton.ca

**Abstract.** The detection of learning styles in adaptive systems provides a way to better assist learners during their training. A popular approach is to fill out a long questionnaire then ask a specialist to analyze the answers and identify learning styles or types accordingly. Since this process is very time-consuming, a number of automatic approaches have been proposed to reduce the number of questions asked. However the length of questionnaire remains an important concern. In this paper, we address this issue by proposing *T-PREDICT*, a novel dynamic electronic questionnaire for psychological type prediction that further reduces the number of questions. Experimental results show that it can eliminate 81% more questions of the Myers-Briggs Type indicators questionnaire than three state-of-the-art approaches, while predicting learning styles without increasing the error rate.

**Keywords:** dynamic adaptive questionnaire, classification, learning styles, Myers-Briggs Type Indicator, psychological types.

## 1 Introduction

Knowing the learning style of learners is important to ensure enhanced and personalized interactions with teachers. Various studies established that presenting information in a manner that is adapted to the student's learning style facilitates learning (e.g. [1, 2]).

Two main approaches are used to carry out learning style or type assessment: the first approach consists of examining the learner's interaction with the system [1, 3]. The second approach entails using a standard questionnaire that is filled out by an individual before a specialist examines the responses and establishes the corresponding learning style.

Even though these approaches can correctly identify the learning type of a learner, they still suffer from various shortcomings. Using the first approach entails that an initial random learning style is assigned to the learner, implying that if the initial guess is wrong, the system will offer wrong interactions with the learner and that could have negative impact on learning. In addition, the interaction will persist until enough data is stored so that ideal learning style can be established [3]. The second approach has various limitations as well. First, the questionnaires tend to be long and time consuming. For instance, Myers-Briggs Type indicator questionnaire has more than 90 questions. Second, the fact that questionnaires are long implies that the person filling the questionnaire might be demotivated to reply to the questions with enough attention [3] which might lead to abandoning the test, skipping questions, answering falsely, etc. Consequently, an incorrect learning style might be adopted [4]. The third limitation of this approach is that in order to assign a learning style, there is a need for a specialist to analyze the learner's answers to questions.

To address these limitations, several approaches [5, 6, 7] have been proposed to automatically reduce the number of questions asked to learners in an effort to identify their learning styles or psychological types. For example, *Q-SELECT* [5] is an adaptive electronic questionnaire that uses association rules to predict part of the answers and hence shorten the questionnaire by up to 30%. However, even when applying these approaches, the length of questionnaires remains a concern. Therefore, an important research question is "could a new method be designed to further reduce the number of questions asked?"

In this paper, we answer this question positively by proposing a dynamic electronic questionnaire *T-PREDICT* that further reduces the number of questions and automatically recognizes the psychological types with high accuracy.

The rest of the paper is organized as follows. The Myers-Briggs Type Indicator model is presented in section 2. Section 3 discusses related work on adaptive questionnaires. Section 4 and 5 respectively present the proposed electronic questionnaire and the experimental results. Finally, section 6 draws the conclusions.

## 2 Myers-Briggs Type Indicator (MBTI)

The Myers-Briggs Type Indicators (MBTI) is a well-known personality assessment model that has been used for over three decades. It uses Carl Jung's personality type theory to categorize individuals into four dimensions. Each dimension consists of two opposite preferences that depict inclinations and reveal dispositions in personal mindsets and techniques of making decisions [8].

The E-I dimension (*extraverted-introverted*) centers its attention on establishing whether a person's approach is influenced by the outward environment such as other objects and individuals, or it is internally-oriented. The S-N dimension seeks to measure *sensing-intuitiveness*, illustrating the perception approach of an individual. As per the T-F dimension, *thinking* implies coherent reasoning and decision making processes while *feelings* influence personal, objective, and value oriented approach. The J-P dimension involves either a *judging* attitude and quick decision-making or *perception* that demonstrates more patience and information gathering prior to decision-making. Given these four dimensions, an individual's personality type is therefore designated by a four-letter code: {ISTJ, ISFJ, INFJ, INTJ, ISTP, ISFP, INFP, INTP, ESTP, ESFP, ENFP, ENTP, ESTJ, ESFJ, ENFJ, ENTJ}.

While a number of the preferences may be dominant, others are likely to be secondary and can be easily subjected by other dimensions. For instance, the J-P dimension influences the two function preferences namely T or F versus S or N [9].

Though the Myers-Briggs Type Indicator questionnaire has been used widely for a long time, it has some limitations. Its hypothetical and numerical imports are restricted to some extent by the use of dichotomous preference items [10]. In addition, the MBTI questionnaire contains numerous questions, a number which may put off some users. As a result, they may choose to answer the questionnaire without paying much thought or attention to the choices thus raising doubts on the reliability of the assessment [11]. Reducing the number of questions in a questionnaire has been a way to increase its efficiency [3, 12].

## 3 Related work

Some major challenges in building an e-learning system that can adapt itself to a learner is giving it the capability of changing the type, the order, or the number of questions presented to the learner [3, 5, 13, 14]. For instance, McSherry [12] reports that reducing the number of questions asked by an informal case-based reasoning system minimized frustration, made learning easier and increased efficiency.

Various methods have been suggested in order to minimize the number of questions that are required to establish the learning style or preference of an individual. The *AH questionnaire* [3] uses decision trees to reduce the questions and categorize the students as per the Felder-Silverman theory of learning styles. From an experiment that had 330 students, it was likely to anticipate the learning styles with precision of up to 95.71% while only asking four or five questions among the eleven questions that are applied in each dimension.

*EDUFORM* is a software used for adaptation and dynamic optimization of questionnaire propositions for online profiling of learners, which was proposed by Nokelainen et al. [6]. The tool, that uses Bayesian modeling as well as abductive reasoning, minimized the questionnaire items by 30% to 50% while retaining an error rate of 10% to 15%. These results showed that shortening questionnaires does not detrimentally influence the correct categorization of individuals.

In previous work [7], two methods based on back-propagation neural networks and decision trees were proposed to predict learning types and reduce the number of questions asked in the MBTI questionnaire. We refer to these methods as *Q-NN* and *Q-DT*, respectively. The general experimental method tries to identify questions that are less influential in determining the learning types. These questions are then eliminated from the questionnaire and the learning types are predicted. This process is repeated until a maximum error rate of 12% is achieved. In an experimental study, that had 1,931 filled questionnaires, the *Q-NN* method clearly identified and eliminated 35% of the questions while establishing the learning preferences with an error rate of 9.4%. On the other hand, *Q-DT* eliminated 30% of the questions with an error rate of 14%.

Recently, we presented an alternative approach to reduce the number of questions in MBTI questionnaire [5]. This approach (*Q-SELECT*) comprises three modules: (1)

an answer prediction algorithm, (2) a dynamic question selection algorithm and (3) an algorithm to accurately predict a person's learning style based on both user supplied and predicted answers. The two first modules rely on association rules between answers from previous users. The third module predicts learning types using a neural network. An experimental study with the same 1,931 MBTI filled questionnaires has shown that *Q-SELECT* reduces the number of questions asked by a median of 30% with an average error rate of 12.1%. The main advantage of *Q-SELECT* compared to *Q-NN* and *Q-DT* is its adaptability, i.e. the ability to ask a variable number of questions and to reorder them depending on the user answers, thus providing a personalized questionnaire to each user.

Reducing the MBTI questionnaire by up to 35% still leaves around 60 questions. This paper presents *T-PREDICT,* a new dynamic approach to further minimize the number of questions in MBTI questionnaires while predicting learning types with a comparable error rate.

Reducing the number of questions might bring to mind algorithms for dimensionality reduction such as PCA, ICA [15]. However, these methods would have reduced the same number of questions for all users as in *Q-NN*. Our choice was to continue our experiments with a dynamic user-adaptive approach. This choice was later confirmed by our results from *T-PREDICT.* We were sometimes able to predict the learning type with as little as six questions out of 92 but with a median of 11 questions.

Another popular theory that deals with questions and answers is *Item Response Theory* (*IRT*) [16]. It has been applied in education and psychology to assess an underlying ability or trait using a questionnaire. To apply *IRT*, users answers need to be collected and analyzed. Using a technique such as factor analysis, a model (e.g. logistic function) is created for each question to represent the amount of information provided by each answer about the latent trait. The quality of the generated models varies depending on the data available. When a model does not fit well a question, this latter is typically removed, replaced or rewritten [17]. Applying *IRT* can be very time-consuming, since for each modification, more data may need to be collected to update models and human intervention is required to analyze questions and tweak models. Furthermore, *IRT* does not provide means for user adaptability. In contrast, our proposal is automatic and user-adaptive in all steps: selecting important questions, reordering them, and later predicting learning types.

## 4 The Electronic Questionnaire

The proposed electronic questionnaire *T-PREDICT* comprises two major components: a question sorting algorithm and a learning type prediction algorithm. The philosophy behind this division is that some questions might be more important than others i.e. their answers could easily classify the learners in one or the other personality type. Hence, the sorting algorithm sorts the questions by their ability to discriminate between classes, while the prediction algorithm uses them in this preferential order. An additional module is developed to dynamically select parameters needed for the prediction algorithm.

### 4.1 The question sorting algorithm

The MBTI questionnaire evaluates each dimension (EI, SN, TF, and JP) by a distinct subset of questions. Thus, we split the questionnaire into the four sets of questions representing these dimensions. Each dimension consists of two opposite preferences, hence the need to classify individuals into one of these two preferences (classes). Since our goal is to reduce the number of questions asked, it is crucial to recognize which questions are more important for identifying the preferences. We define the importance of a question as its ability to discriminate between the two classes. Let $q$ be a question and $A(q) = \{a_{1 \ldots} a_m\}$ be the set of possible answers to this question. Let $T$ be a training set of filled questionnaires such that each questionnaire belongs to one of the two classes. For any answer $a_i$ ($1 \leq i \leq m$), let $N_1(a_i)$ and $N_2(a_i)$ respectively be the number of questionnaires in $T$ containing the answer $a_i$ that belong to the first class and the second class. The *discriminative power* of the answer $a_i$ is denoted as $DA(a_i)$ and defined as $N_1(a_i) / N_2(a_i)$ if $N_1(a_i) > N_2(a_i)$ or $N_2(a_i) / N_1(a_i)$, otherwise. Intuitively, it represents how many times one class is larger than the other for individuals having answered $a_i$, and thus how this answer helps to discriminate between the two classes. The *discriminative power* of a question $q$ is denoted as $DQ(q)$ and defined as $DQ(q) = \sum_{k=1}^{|A(q)|} DA(a_i) / |T|$. The proposed adaptive questionnaire initially calculates the discriminative power of each question in the training set $T$ and sorts them accordingly (see Fig. 1 for the pseudocode). The next subsection describes how the adaptive questionnaire asks the questions by decreasing order of discriminative power and predicts the preference (class) of each individual using the provided answers.

**QUESTION_SORT** (a training set $T$, a list of questions $Q$, a set of possible answers $A$)

**FOR** each question $q \in Q$
    **FOR** each possible answer $a_i \in A(q)$
        **SCAN** $T$ to calculate $N_1(a_i)$ and $N_2(a_i)$.
        **IF** $N_1(a_i) > N_2(a_i)$ **THEN** $DA(a_i) := N_1(a_i) / N_2(a_i)$.
            **ELSE** $DA(a_i) := N_2(a_i) / N_1(a_i)$.
    **END FOR**
    $DQ(q) = \sum_{k=1}^{|A(q)|} DA(a_i) / |T|$.
**END FOR**
**SORT** $Q$ such that $q_a$ appears before $q_b$ if $DQ(q_a) > DQ(q_b)$, for all questions $q_a, q_b \in Q$.
**RETURN** $Q$.

**Fig. 1.** The question sorting algorithm

### 4.2 The learning style prediction algorithm

The learning style prediction algorithm (see Fig. 2) automatically identifies the preference of a user in a given dimension based on supplied answers and their similarity to answers from previous users. The algorithm takes as input a training set of filled questionnaires $T$, the list of questions $Q$ for the dimension, sorted by their discriminative power (see section 4.1), a maximum number of questions to be asked *maxQues-*

*tions* (by default set to $|Q|$), and two additional parameters *minMargin* and *minQ* that will be defined later.

---

**PREDICT** (a training set $T$, a list of questions $Q$ sorted by their discriminative power, a maximum number of questions to be asked *maxQuestions*, a minimum margin between classes *minMargin*, a minimum number of questionnaires to be matched *minQ*)

$U := \emptyset$.<br>
**FOR** each question $q_i \in Q$ until *maxQuestions* have been asked<br>
    **ASK** $q_i$ to the user and store the provided answer in $U$.<br>
    $S_1 := S_2 := 0$.<br>
    **FOR** each questionnaire $X \in T$ such that $U \subseteq X$<br>
        **IF** $X$ is tagged as class 1 **THEN** $S_1 := S_1 + 1$. **ELSE** $S_2 := S_2 + 1$.<br>
    **IF** $S_1 + S_2 \geq minQ$ **AND** $|S_1 - S_2| \geq minMargin$ **THEN**<br>
        **IF** $S_1 > S_2$ **THEN RETURN** *"class 1"*. **ELSE RETURN** *"class 2"*.<br>
**END FOR**<br>
$Y_1 := Y_2 := 0$.<br>
**FOR** each questionnaire $X \in T$ such that $C(X, U) \geq |U|/2$<br>
        **IF** $X$ is tagged as class 1 **THEN** $Y_1 := Y_1 + C(X, U)$. **ELSE** $Y_2 := Y_2 + C(X, U)$.<br>
**IF** $Y_1 > Y_2$ **THEN RETURN** "*class 1*". **ELSE RETURN** "*class 2*".

---

**Fig. 2.** The learning type prediction algorithm

The algorithm first initializes a set $U$ as empty to store the user answers. Then, a loop is performed to ask the questions to the user in decreasing order of their discriminative power. Each provided answer is immediately added to $U$. Then, the algorithm attempts to make a prediction by scanning through the set $T$ to count how many users have the same answers as the current user and belong to class 1 or belong to class 2, i.e. $S_1 = |\{X \mid X \in T \wedge U \subseteq X \wedge X \text{ is tagged as class } 1\}|$ and $S_2 = |\{X \mid X \in T \wedge U \subseteq X \wedge X \text{ is tagged as class } 2\}|$. Two criteria must be met to make this prediction. First, the number of filled questionnaires, from the training set $T$, matching the answers of the current user should be higher or equal to a pre-set threshold (*minQ*), i.e. $S_1 + S_2 \geq minQ$. Second, there should be a large enough difference between $S_1$ and $S_2$ in order to make an accurate prediction. This difference is defined as $|S_1 - S_2| \geq minMargin$, where *minMargin* is also a pre-set threshold. If both conditions are met, a prediction is made. The prediction is class 1 if $S_1 > S_2$, and class 2 otherwise. If no prediction can be made, the algorithm continues with other questions, one at a time, until it is able to make a prediction.

After *maxQuestions* questions have been exhausted and no prediction was possible with the exact matching, a prediction is made by considering an *approximate* match between the present user answers and questionnaires from the training set $T$. A questionnaire $X \in T$ approximately matches $U$ if it shares at least $|U|/2$ answers with it. The number of answers common to $X$ and $U$ is denoted as $C(X, U)$ and defined as $C(X, U) = |X \cap U|$. Let $Y_1$ and $Y_2$ be the sets of all questionnaires from $T$ that approximately match $U$ and are tagged as class 1 or class 2. The prediction is class 1 if $Y_1 > Y_2$. Otherwise, it is class 2.

### 4.3 The parameters selection algorithm

As mentioned earlier, the prediction algorithm needs two preselected parameters *minMargin* and *minQ*. These parameters are not global. They are dynamically selected as per number of questions asked. The *PARAMETERS_SELECT* algorithm (see Fig. 3) attempts to select the best values for these parameters exhaustively by simulating predictions and calculating the corresponding precisions on a test set of filled questionnaires *TS* by using a training set *T*.

```
PARAMETERS_SELECT (a training set T, a test set TS, a list of questions Q= {q1, q2,…qn}
sorted by their discriminative power, a target precision TargetPrecision, and the two upper
bounds maxMargin and maxQ)
1.  GlobalPrecision :=0.
2.  RequiredPrecision := TargetPrecision
3.  P := ∅.
4.  RemainingQuestionnaires := TS.
5.  FOR each qk ∈ Q and k := 1, 2…n
6.      bestScore := 0.
7.      bestParameters := (0, 0, 0).
8.      bestResult := (0, 0, 0).
9.      FOR each combination of values i, j, α such that i :=1, 2… maxQ
            and j := 1, 5,… maxMargin
            and α := 0.5, 0.6 … 0.9
10.         Predict a class for each questionnaire X ∈ RemainingQuestionnaires using
            the first k questions and the training set T.
            Set parameters minQ := i, minMargin := j for the kth question, while keeping the
            best previously found parameters for questions 1 to k-1.
            Let predictedk be the set of questionnaires where a prediction was performed
            and precisionk be the precision.
11.         score := precisionk × α + |predictedk| × (1- α).
12.         IF score > bestScore AND precisionk ≥ RequiredPrecision THEN
13.             bestParameters := (k, i, j).
14.             bestScore := score.
15.             bestResult := (k, predictedk, precisionk).
16.         END IF
17.     END FOR
18.     P: = P ∪ {bestParameters}.
19.     GlobalPrecision := GlobalPrecision +|bestResult.predicted| × bestResult.precision
20.     RemainingQuestionnaires := RemainingQuestionnaires \ bestResult.predicted.
21.     RequiredPrecision := (|TS|×TargetPrecision – GlobalPrecision )
                                / RemainingQuestionnaires
22. END FOR
23. RETURN P.
```

**Fig. 3.** The parameter selection algorithm

The algorithm proceeds by considering each question from $Q$ in the order that they will be asked by the PREDICT algorithm (in descending order of discriminative power). For every question $q_k$ considered, all combinations of values *i* and *j* for parameters *minQ* and *minMargin* in their respective intervals [1, *maxQ*] and [1, *maxMargin*] are tested. For each combination, a certain number of predictions $NbPredictions_k$ are obtained with a corresponding precision $precision_k$. The combination of values $i$ and $j$ that is considered the best is the one maximising the function *score* $(i,j) = \alpha \times precision_k + NbPredictions_k \times (1- \alpha)$, where α is a constant varied in the [0.5, 0.9] interval. The weight α is used to calibrate the relative importance of the precision and the number of predictions on the selection of parameter values.

Furthermore, an additional constraint is that the best combination needs to have a precision higher than a moving threshold *RequiredPrecision*. For the first question, this threshold is equal to a pre-set minimum precision *TargetPrecision*. For any subsequent question $q_k$, the required precision is recalculated to take into account the precisions obtained with previous questions (called $GlobalPrecision_k$). The reason is that if a high precision is obtained for the previous questions, it is possible to accept a lower precision for the next questions, while maintaining a global precision above *TargetPrecision*. The global precision is calculated by weighting the previous precisions with their corresponding number of predictions. Formally, the global precision is defined as: $GlobalPrecision_k = \sum_{f=1}^{k-1} precision_f \times NbPredictions_f)$. The required precision for the $k^{th}$ question is calculated as: $RequiredPrecision_k = (|TS| \times Target\ Precision - GlobalPrecision_k) / |RemainingQ|$, where *RemainingQ* is the number of unpredicted questionnaires.

When the algorithm terminates, it returns a list $P$ of triples of the form ($k$, $y$, z) indicating that parameters *minMargin* and *minQ* should be respectively set to values $y$ and $z$ for the $k^{\text{th}}$ question to establish a prediction with a global precision not below *TargetPrecision*.

## 5 Experimental Results

Two experiments were performed. The first one compares the performance of the learning type prediction algorithm *T-PREDICT* with other methods. The second one assesses the influence of question sorting and limiting the number of questions in *T-PREDICT*. A database of 1,931 MBTI filled out questionnaires were supplied for experimentation by Prof. Robert Baudouin, an expert of the MBTI technique at Université de Moncton. The number of samples used for training was 1,000 while 931 were used for testing. The proposed dynamic questionnaire *T-PREDICT* is implemented using Java. The goal of the experiments was to ask as few questions as possible while automatically identifying learning types with a minimum precision of 88 %, i.e. a maximum error rate of 12%. This is the same error rate that was used in *Q-SELECT* [5], thus allowing a more precise comparison base.

Parameters used were automatically selected by the *parameters selection algorithm* (see section 4.4). It varied *minQ* and *minMargin*, for each question, in the [0.01, 0.15] and [0.1, 0.9] intervals respectively and returned their optimal values to be used by the *type prediction algorithm*. These predictions were done for each of the four MBTI dimensions.

### 5.1 Comparison with other methods

The first experiment compared the median number of questions asked in each dimension with results obtained by three specialized methods that have been specifically developed to reduce the number of questions for MBTI. These methods are based on back-propagation neural networks (*Q-NN*) [7], decision trees (*Q-DT*) [7], and association rules (*Q-SELECT*) [5]. Table 1 shows the number of questions asked per dimension, as compared to the original higher number, and the corresponding error rates. Note that the *T-PREDICT* line shows that the number of questions asked for EI, SN, TF and JP dimensions were 2 out of 21 questions, 4 out of 25 questions, 2 out of 23 questions, and 3 out of 23 questions respectively. The corresponding error rates obtained were 10.7% for EI, 13% for SN, 10.3% for TF and 13.1% for JP. Thus, maintaining a weighted average error rate of 12.1% for the four dimensions, which is practically the pre-set error rate of 12% mentioned earlier. All dimensions for the four compared methods maintained error rates below 13.1% except the TF and JP dimensions for the *Q-DT*, which respectively had error rates of 16.3% and 13.7%. Overall, *T-PREDICT* asked a median of only 11 questions, out of 92 questions, to predict learning types with an average error rate of 12.1%, while *Q-SELECT*, *Q-NN*, and *Q-DT* asked 62, 60, and 64 questions to achieve error rates of 11.4%, 9.4%, and 14% respectively. In sum, the number of questions was greatly reduced by *T-PREDICT* while maintaining a very comparable error rate.

**Table 1.** Median number of questions asked by *T-PREDICT* per dimension, compared to other methods.

| | EI (21 quest.) | SN (25 quest.) | TF (23 quest.) | JP (23 quest.) | Total (92 quest.) |
|---|---|---|---|---|---|
| *T-PREDICT* | **2** | **4** | **2** | **3** | **11** |
| Avg. error rate | 10.7% | 13% | 10.3% | 13.1% | 12.1% |
| *Q-SELECT* | 14 | 16 | 14 | 18 | 62 |
| Avg. error rate | 9.9% | 11.8% | 12% | 11.7% | 11.4% |
| *Q-NN* | 14 | 16 | 16 | 14 | 60 |
| Avg. error rate | 8.3% | 10.9% | 9.3% | 8.9% | 9.4% |
| *Q-DT* | 14 | 17 | 17 | 16 | 64 |
| Avg. error rate | 12.8% | 13.1% | 16.3% | 13.7% | 14% |

Since the above table presents median numbers of questions asked, one might wonder how the actual numbers of questions asked are distributed. For example, Table 2 shows the distribution of questions asked for the SN dimension. Note that 34% of individual learning types were predicted using just one question, a cumulative 76% were predicted using four questions, and the remaining 24% were predicted using more questions. All 100% of the predictions were possible with at most thirteen questions. However, the error rate increases to 25% when thirteen questions are used. This is due to the fact that after the maximum number of questions (*maxQuestions*) has been exhausted, predictions are made by approximate matching.

**Table 2.** Distribution of questions asked by *T-PREDICT* for the SN dimension

| Number of questions asked $m$ | Number of learning types predicted using $m$ questions (in %) | Error rate for predictions using $m$ questions | Cumulative percentage of learning types predicted using up to $m$ questions |
|---|---|---|---|
| 1 | 313 (34%) | 4.8% | 34% |
| 2 | 0 (0%) | - | 34% |
| 3 | 96 (10%) | 5.2% | 44% |
| 4 | 298 (32%) | 17.1% | 76% |
| 5 | 0 (0%) | - | 76% |
| 6 | 0 (0%) | - | 76% |
| 7 | 0 (0%) | - | 76% |
| 8 | 47 (5%) | 14.9% | 81% |
| 9 | 0 (0%) | - | 81% |
| 10 | 0 (0%) | - | 81% |
| 11 | 0 (0%) | - | 81% |
| 12 | 0 (0%) | - | 81% |
| 13 | 0 (0%) | - | 81% |
| 13 (approx. match) | 177 (19%) | 25% | 100% |

Note that some questions, such as the second question in Table 2, did not generate any additional predictions. These are the questions where the two established criteria necessary to make a prediction, as set by the *prediction algorithm*, were not met.

### 5.2 Influence of sorting and limiting number of questions

The following experiment assesses the influence of two optimizing strategies on error rate and median number of questions asked, these are (1) sorting questions by discriminative power and (2) limiting the maximum number of questions asked. *T-PREDICT* results were compared with a modified version that does not sort questions (*UT-PREDICT*). Moreover, the maximum number of questions asked (*MaxQuestions*) was varied from one to all questions. Table 3 shows the obtained results for the SN dimension. Results for other dimensions were similar, although not shown due to space limitation. It can be observed that the median number of questions is higher and that the error rate is much higher when the questions are unsorted. For example, Table 3 shows that *T-PREDICT* had an error rate of 13% using a median of 4 questions, while *UT-PREDICT* used 5 questions with an error rate of 17.84%, when *MaxQuestions* = 13. It can also be noticed that limiting the maximum number of questions to 13 gave the lowest error rate (13 %).

**Table 3.** Influence of *MaxQuestions* and sorting questions by discriminative power on error rate and median number of questions asked

| Parameter *MaxQuestions* | Median number of questions (*T-PREDICT*) | Average error rate (*T-PREDICT*) | Median number of questions (*UT-PREDICT*) | Average error rate (*UT-PREDICT*) |
|---|---|---|---|---|
| 1 | 1 | 32,8% | 1 | 32,8% |
| 2 | 2 | 32,8% | 2 | 32,8% |
| 3 | 3 | 18,1% | 3 | 28,0% |
| 4 | 4 | 18,1% | 4 | 32,3% |
| 5 | 4 | 15,4% | 5 | 20,8% |
| 6 | 4 | 18,4% | 5 | 22,6 % |
| 7 | 4 | 16,3% | 5 | 20,0% |
| 8 | 4 | 14,7% | 5 | 21,4% |
| 9 | 4 | 14,5% | 5 | 17,9% |
| 10 | 4 | 15,0% | 5 | 19,2% |
| 11 | 4 | 14,0 % | 5 | 17,9% |
| 12 | 4 | 13,7% | 5 | 19,2% |
| **13** | **4** | **13,0**% | **5** | **17,8**% |
| 14 | 4 | 13.9% | 5 | 19.6% |
| 25 | 4 | 15.1% | 5 | 18.3% |

## 6 Conclusion

Filling questionnaires for learning style assessment is a very time-consuming task, which might lead to abandoning the test, skipping questions, answering falsely, etc. To address this issue, various approaches have been proposed to reduce the size of questionnaires. In this paper, we have presented a novel dynamic electronic questionnaire *T-PREDICT* to further reduce the number of questions needed for learning type identification. It comprises three modules: a question sorting algorithm, a prediction algorithm and an automatic parameter selection algorithm.

Experimental results with 1,931 filled questionnaires for the Myers Briggs Type Indicators show that our novel approach asked a median of only 11 out of 92 questions to predict learning types, with an average error rate of 12.1%, while previous approaches *Q-SELECT* [5], *Q-NN* and *Q-DT* [7] asked between 60 and 64 questions to achieve error rates between 9.4% and 14%. Another defining characteristic of *T-PREDICT* is its ability to ask a variable number of questions, thus providing an automatic personalized questionnaire to each user, like *Q-SELECT* but unlike *Q-NN*, *Q-DT*, *PCA* and *IRT* that apply the same reduction to all users.

## References


1. Felder, R. M., Brent, R.: Understanding student differences. J. of Engineering Education. 94(1), 57–72 (2005)
2. Graf, S.: Adaptivity in Learning Management Systems Focusing on Learning Styles. Ph.D. Thesis, Vienna University of Technology, Vienna (2007)
3. Ortigosa, A., Paredes, P., Rodríguez, P.: AH-questionnaire: An adaptive hierarchical questionnaire for learning styles. Computers & Education. 54(4), 999–1005 (2010)
4. García, P., Amandi, A., Schiaffino, S., Campo, M.: Evaluating Bayesian Networks' Precision for Detecting Students' Learning Styles. Computers & Edu. 49 (3), 794–808. (2007)
5. Mwamikazi, E., Fournier-Viger, P., Moghrabi, C., Barhoumi, A., Baudouin, B.: An Adaptive Questionnaire for Automatic Identification of Learning Styles. In: Proc. IEA AIE 2014. Springer, Heidelberg, pp. 309-409 (2014)
6. Nokelainen, P. et al.: Bayesian Modeling Approach to Implement an Adaptive Questionnaire. In: Proc. ED-MEDIA 2001, pp. 1412–1413, AACE, Chesapeake (2001)
7. Barhoumi, A.: Simplification du questionnaire MBTI par apprentissage automatique en vue de faciliter l'adaptabilité des logiciels de formation en ligne. M.Sc. Thesis. Université de Moncton, 117 pages (2012)
8. Boyle, G. J.: Myers-Briggs Type Indicator (MBTI): Some psychometric limitations. Australian Psychologist. 30(1), 71–74 (2009)
9. El Bachari, E., Abdelwahed, E. H., EI Adnani, M.: Design of an Adaptive E-Learning Model Based on Learner's Personality. Ubiquitous Comp. Comm. J. 5(3), 27–36 (2010)
10. Francis, L. J., Jones, S. H.: The Relationship between MBTI and the Eysenck Personality Questionnaire among Adult Churchgoers. Pastoral Psychology. 48(5), 377–386 (2008)
11. Pittenger, D., J.: Measuring the MBTI and coming up short. Journal of Career Planning and Employment. 54(1), 48–52 (1993)
12. McSherry, D.: Increasing dialogue efficiency in case-based reasoning without loss of solution quality. In: Proc. IJCAI 2003, pp. 121–126. Morgan Kaufmann (2003)
13. Abernethy, J., Evgeniou, T., Vert, J.-P.: An optimization framework for adaptive questionnaire design. Technical Report, INSEAD, Fontainebleau, France (2004)
14. Chen, C.-M., Lee, H. M., Chen, Y.-H.: Personalized e-learning system using item response theory. Computers & Education. 44(3), 237-255 (2005)
15. Jolliffe I.T.: Principal Component Analysis 2nd ed. Springer, New York (2002)
16. DeMars, C.: Item response theory. Oxford University Press, Oxford (2010)
17. Weinhardt, J. M., Morse, B. J., Chimeli, J., Fisher, J.: An item response theory and factor analytic examination of two prominent maximizing tendency scales. Judgment & Decision Making, 7(5), 644–658 (2012)