

Galli, Richard

**Conference Paper**

## Using quality information is analysis - with experiment

22nd European Regional Conference of the International Telecommunications Society (ITS):  
"Innovative ICT Applications - Emerging Regulatory, Economic and Policy Issues", Budapest,  
Hungary, 18th-21st September, 2011

**Provided in Cooperation with:**

International Telecommunications Society (ITS)

*Suggested Citation:* Galli, Richard (2011) : Using quality information is analysis - with experiment, 22nd European Regional Conference of the International Telecommunications Society (ITS): "Innovative ICT Applications - Emerging Regulatory, Economic and Policy Issues", Budapest, Hungary, 18th-21st September, 2011, International Telecommunications Society (ITS), Calgary

This Version is available at:

<https://hdl.handle.net/10419/52174>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

**Richard GALLI**

**Using quality information in analysis – with experiment**

**Abstract:** Using quality information for analysis, or staying with just quantity based information is always a good base for debate: Either relying on quantitative measurements and data (which is sad to be more accurate and easier to rely on) or use qualitative measurements (which is sad to be human / easier to understand and answer). In both ways the collected data will require further processing to be used in decision-making.

An experiment was made on customer satisfaction on an imaginary service. The service was a web based TTS (Text-To-Speech) software test. Of course, all the voices were pre-recorded, so the participants were all chosen not to realize that (every member with low IT knowledge). After the end of the reading everyone had to answer a few questions regarding the quality. Group “A” had to answer the questions (which were sounding more scientific) with numbers related the current measure – help provided if needed. Group “B” had to answer the same questions with open answers, as they like.

**Keywords:** quality data, quality measures, qualitative and quantitative measurements for CRM/BI

**Authors' affiliation and corresponding author's e-mail address**

**Richard GALLI, assistant**

**Széchenyi István University, Department of Information Sciences**

**H-9026 Győr, Egyetem tér 1.**

**richard@sze.hu**

There are many situations, when the same incidence can be expressed qualitatively and quantitatively. In situations, when the domain (range) and the borders are available – quantitative data can be used very well. But in many cases, when defining borders is hard in itself, or when classification is dependent on other factors, could be better to rely on qualitative data. One of the reasons is, that the measured value is hard to give a dimension, or it is hard to describe it with numbers. On the other hand qualitative attributes are much easier for the human brain to understand. (The best example may be the user satisfactory measurements)

As for the qualitative information: it is thought to be richer and well grounded, even in those cases, when there is one-to-one matching between qualitative and quantitative information.

When asking on qualitative data, the openness of the question is also important. When closed questions are applied, the respondent could feel pressed – on the other hand it is trivial, if the question is well interpretable and the dimensions are present (if required) the fill out can be much faster, and the results more reliable.

Answers given to open questions better reflect the respondent's way of thinking– likewise, to the analysis, creating a rule system is required, which may take considerable time.

The goal of a qualitative analysis is not only describing the data, but describing those objects or events which resulted the data. Of course, in most cases, description is not enough, those events have to be interpreted, explained, or forecasted (relying on data).

The three key components of qualitative data analysis are the data description, their classification and then making relations between them.

The examined phenomenon the description of a hypotenuse is called thick description, since he implies extra information from the environment, the aims and from that direction, that in case of given outcome what kind of act can be expected him. Forecasting a variable measured on a discreet category scale is the classification or classification. To the production of the result several explanatory ones (access) it is necessary to examine his changing combination, then they to assign the output based on his values.

The comprehensive description of the examined phenomenon is called thick description, because it includes some extra information about the environment, the goals and actions

depended on the various outcomes. Classification is the marking of the measured variable on a discreet category scale. To reach results, several explanatory variables – and their combinations – must be analyzed, and then according to these values assign the output.

All classification problems can be explained the same way: a teaching set can be given, with preassigned input-expected output values. The task is to create a rule, which can classify every input value

It is easier to perceive the classification task as an estimation, and estimate the conditional probability for the input data belonging to a given classification class.

The other emerging problem is whether binary or multiclass classification should be applied. Binary classification means only two categories are present, by multiclass cases can be more. Since for binary classification many methods were developed, it is easier to trace back every multiclass-required case to more binary classifications.

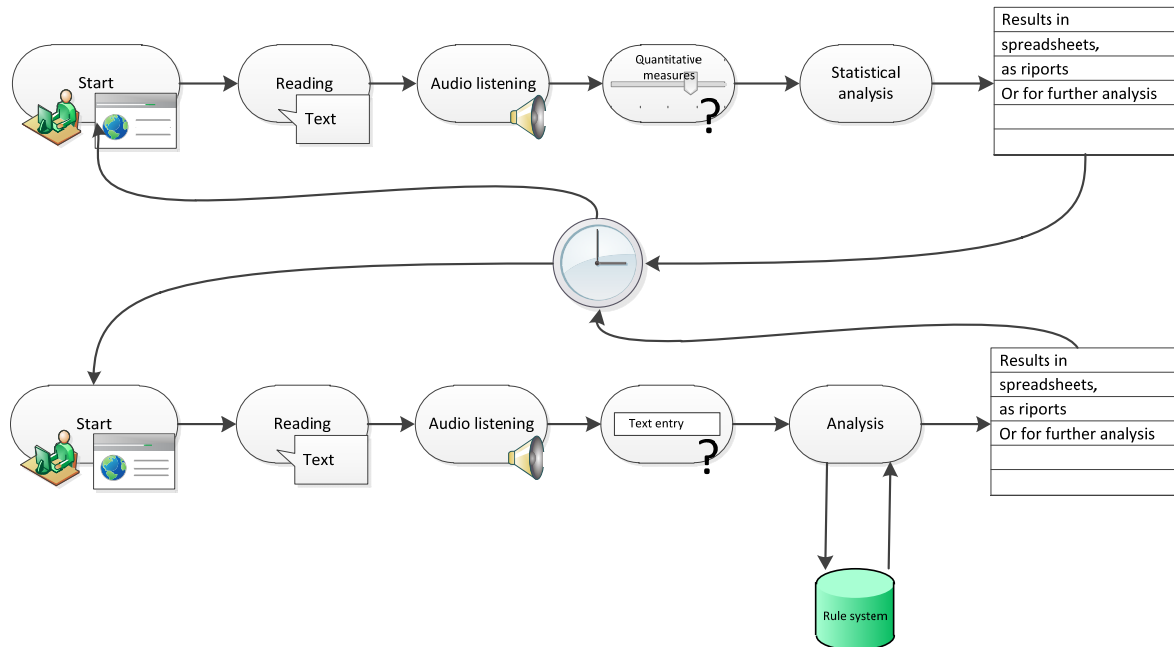
#### Using quality data; experiment for satisfaction measurement:

The initial situation: Every user sits in front of a computer (singly), then they see some plain text, and then a software reads out the text for them. This is a fictitious, simulated service. The fictitious product is a new TTS (text To Speech) system. According to the experiment our goal is to measure the service levels and user satisfaction.

The script is the following: The user enters, the site asks for an e-mail address (for later identification). Every user sees the same text, they cannot change this. Then onto pressing a button the system reads out the text apparently continuously – as if the processing would happen locally. (Only those were selected for the experiment, who are strictly no IT experts.) Then the user listens to the reading. The system has more soundtracks, each is different in quality (the worst one is nearly impossible to understand, the best is with a perfect pronunciation; naturally not every one was made with a real TTS software, there are human sounds between them). After the listening the quality of the service must be evaluated. A group of the users must do this in a quantitative way, the other group must do this part based on qualitative criteria. (The system does not make a choice randomly between the readings, it is predefined, which user hears which sound.) After filling in the form (evaluation), the users are asked to come back in two weeks and repeat the measurement. After the break the experiment is reversed, who did qualitative measurement the first time, does quantitative now, and vice versa. To avoid the assumption, that the qualitative attributes are change to a

numerical values in the background, all the qualitative information have to be granted in the form of an open answer.

The algorithm for the experiment:



The questions:

1. Was the reading natural (humanlike)?
2. Was the speed of the reading adequate?
3. Was the rhythm of the reading appropriate to its meaning?
4. Was the tone of the reading adequate?
5. Was the tone of the reading appropriate to its meaning?
6. Was the intonation of the reading lifelike?
7. Was the intonation of the reading appropriate to its meaning?
8. Were the word connections lifelike?
9. Were the syllable connections lifelike?
10. Was the intonation of the sentences appropriate to its meaning?
11. Was the used volume adequate?

The reason for these questions is, that the quality of the mentioned service is hard to quantify (it is possible though, but hard to understand).

A 15 second answer time could be accounted as normal, because this way filling in the whole questionnaire takes 90 seconds, which is acceptable for satisfactory analysis.

Whoever fills it in under significantly shorter time, does not interpret the questions or does not take it seriously, and the results will be useless. On the other hand, who requires much more time to fill it in, probably would not do the same in a real life situation, would just close the form.

The input data and expected outcome data of the experiment:

ID/Soundtrack	Description	Criteria: Quality	Naturality		Speed		Rhythm		Tone		Tone-meaning		Intonation		Intonation-meaning		Word connections		Syllable connections		Sentence-emphasis		Volume	
			1	1	2	2	3	3	4	4	5	5	6	6	7	7	8	8	9	9	10	10	11	11
Audio 1		Best	10	Y	20 s	OK	165 wpm	OK	0	OK	10	OK	10	OK	10	OK	10	OK	10	OK	10	OK	65	OK
Audio 2			9		24 s	OK	147 wpm	OK	2	OK	9	OK	9	OK	10	OK	9	OK	10	OK	9	Bad	70	OK
Audio 3			8		18 s	OK	180 wpm	OK	1	OK	9	OK	7	Bad	10	OK	9	OK	10	OK	9	OK	77	Bad
Audio 4			7		22 s	OK	154 wpm	OK	0	Bad	8	OK	8	OK	7	Bad	8	OK	8	OK	8	Bad	60	OK
Audio 5			6		25 s	Bad	136 wpm	Bad	10	Bad	3	Bad	5	Bad	6	Bad	5	Bad	9	OK	3	Bad	56	Bad
Audio 6			5		19 s	OK	159 wpm	Bad	-4	Bad	3	OK	4	Bad	5	Bad	6	OK	8	OK	3	Bad	70	OK
Audio 7			4		27 s	Bad	133 wpm	OK	-14	Bad	2	Bad	2	Bad	2	Bad	4	Bad	9	Bad	2	Bad	82	Bad
Audio 8			3		20 s	OK	160 wpm	OK	4	Bad	1	Bad	1	Bad	1	Bad	4	Bad	3	Bad	1	Bad	44	Bad
Audio 9			2		28 s	Bad	128 wpm	Bad	-25	Bad	1	Bad	1	Bad	1	Bad	3	Bad	3	Bad	1	Bad	50	Bad
Audio 10		Worst	1	N	15 s	Bad	198 wpm	OK	30	Bad	1	Bad	1	Bad	1	Bad	4	Bad	1	Bad	1	Bad	65	OK

The results:

To review the results we call group “A” (quantitative-qualitative) the group, who had to fill in the satisfactory form based on quantitative criteria first, and then, after the pause based on qualitative criteria. So, group “B” (qualitative-quantitative) – those who had to evaluate the reading based on qualitative criteria first, and in the second turn had to use quantitative answers.

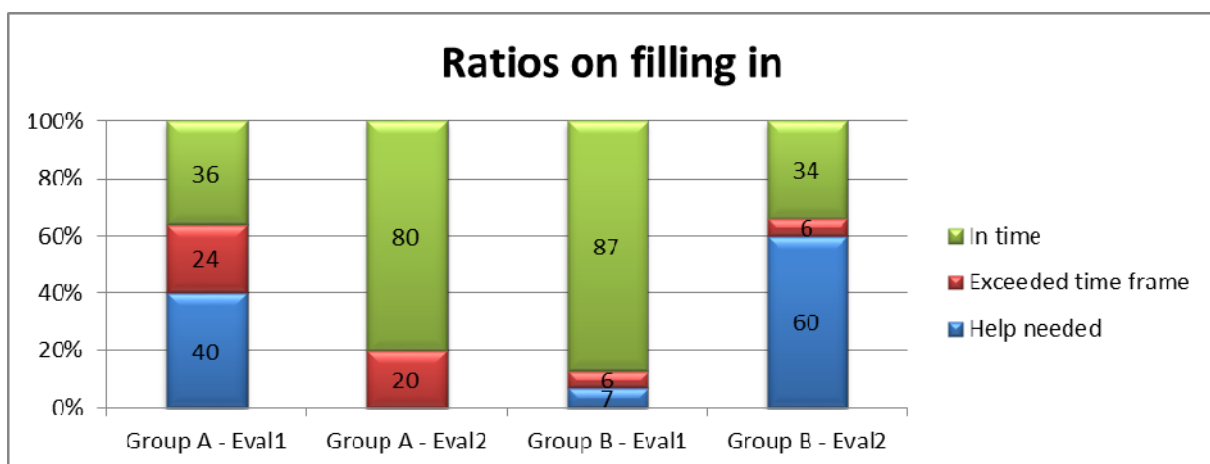
Group “A” results: the first round of the service evaluation met difficulties. 40 percent of the focus group needed continuous help during the evaluation (the reason was not computer usage difficulty, but interpreting the questions and understanding dimensions). 36 percent of the group needed significantly more time to fill the questionnaire (2-2.5 times the time

expenditure) and the vast majority declared in a real life would not fill in the questionnaire, or would not fill it in with the proper care (90%). The remaining 24% filled it in under the expected time frame.

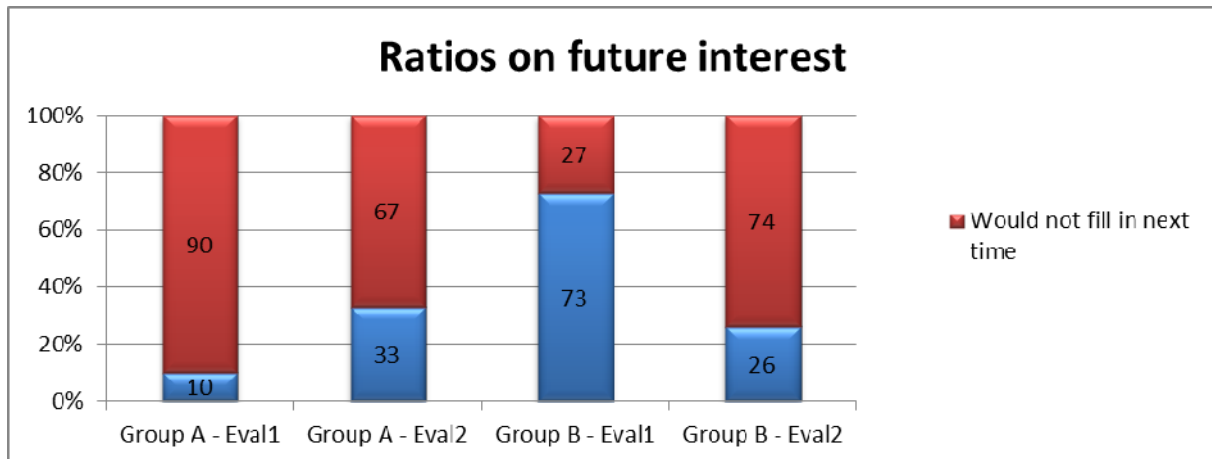
Results for group “A”, after the break are the following: Compared to the previous round, the results are much better, only 20% required more time, than the precalculated time frame, no one required help to fill in the questionnaire (one reason can be for this, that they were already familiar with the occurring concepts). So, 80% of the group was able to fill it in under the target time frame. 67 percent of the group declared, that he/she would not fill in the form in a real life situation.

Group “A” results: the first round of the service evaluation went significantly smoother; only 7% of the group members required explanatory help(the was the problem with the questions interpretation). 87% of the group was able to fill it in under the target time frame. The remaining 6% did it without trouble, but they filled in the questionnaire exceeding the time frame (1.5 times the time expenditure). 73 percent of the group declared, that he/she would fill in the form in a real life situation.

Results for group “A”, after the break are the following: 60% of the group needed continuous help (the reason was never computer usage difficulty, but interpreting the questions and understanding dimensions). 34% of the group was able to fill it in without help, but finished beyond the recommended time frame. The remaining 6% solved the problem in the recommended time frame. 26 percent of the group declared, that he/she would fill in the form in a real life situation.

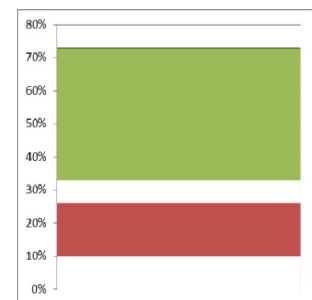


The reason for the second round was to filter out the factors deforming the results due to the groups' composition.



Assessment, inferences and comments:

- If quantitative criteria it to be answered, between 6% and 24% is to be expected the successful filling in of the questionnaire.
- Conversely, applying qualitative criteria, then between 80% and 87% is to be expected the successful filling in.
- In case of quantitative criteria, between 10% and 26% is the filling in expected, if it is presented as an important questionnaire.
- In case of qualitative criteria, between 33% and 73% is the filling in expected in a real life situation, if it is presented as an important questionnaire.
- The usage of quantitative criteria could increase the organizational time of the filling in since it is necessary to provide help in this case.
- By insignificant levels, but the rule-based evaluation takes less time. (But – especially with small samples – the creation of the rule system takes much more time, then spared.)



### Bibliography:

1. Ian Dey: Qualitative Data Analysis: A User-Friendly Guide for Social Scientists. Routledge, 1993.
2. Xindong Wu – Vipin Kumar: The Top Ten Algorithms in Data Mining. CRC Press, 2009.



3. Fabrice Guillet – Howard J. Hamilton: Quality Measures in Data Mining. Springer, 2007.
4. Stanislaw Kozielski – Robert Wrembel: New Trend sin Data Warehousing and Data Analysis. Springer, 2009.