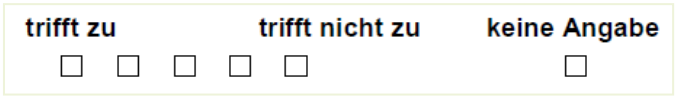
INTERPRET PDF REPORT. BUT HOW?

In the past, questions have repeatedly arisen about the PDF reports that are automatically generated by evasys. In particular, these queries concerned the graphics that are created to evaluate scale questions. The technical statistical terms described in these graphics are explained in this manual.

THE QUESTION TYPE USED: SCALE QUESTIONS

Scale questions are questions in which a fact, a statement or a question is evaluated. In most cases, five checkboxes and one abstention box are displayed. In specific cases, however, the number of checkboxes may vary or an option to abstain may not be given. These are two examples of scale questions as they are currently used at RUB:



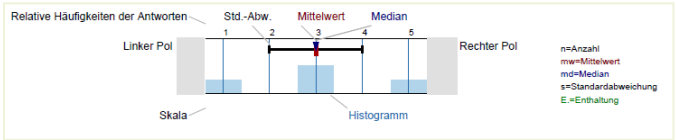
ILLUSTRATIONS: BAR CHART & PROFILE LINE

Scale questions are presented in two forms. The first form corresponds **to a bar chart** in which further statistical measures are graphically incorporated. In addition to the diagram, these measures are also displayed numerically, i.e. as a number.

The second form is the **profile line**, in which the mean values of all scale questions are displayed one above the other in an overview.

The more complex bar chart with its measures is described below and the profile line is presented later.

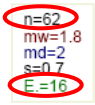
BAR CHART



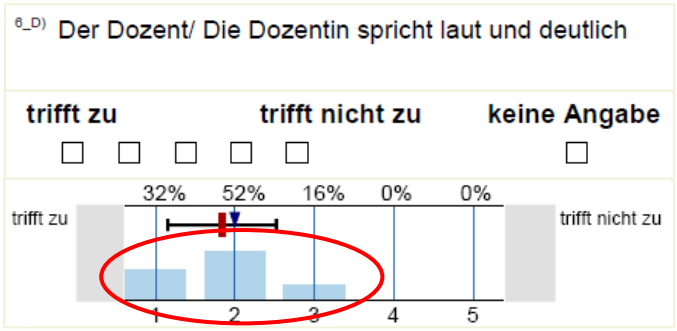
Each PDF report opens with a legend for the bar chart and an overview of the measurement figures. The information to be taken from this diagram is explained below.

QUANTITIY

The list next to the diagram provides an overview of some important measures. **"n"** denotes the number of students who answered the respective question (technical term "base"). In the example shown below for the question "The lecturer speaks loudly and clearly." 62 students responded to this question. Another number is hidden behind the abbreviation **"E"**. In the example, E=16 means that 16 students ticked the box "no response" or "no opinion". If the questionnaire does not include an option to abstain, these are not shown.



The bar chart (see below) clearly shows how the responses (in this example n=62) are distributed across the individual response options. The example shows that 32% of the respondents ticked the first box on the scale ("applies"). 52% chose the second box, 16% the middle category between "applies" and "does not apply".



Whether the example represents a positive or negative result depends on the question or statement that was assessed here. In the example, it is the statement "The lecturer speaks loudly and clearly". It is therefore a positive formulation. The results in the example show that this statement is agreed with to varying degrees. No one rejects this statement. This tends to be a good result.

However, the statement in the example could also have been formulated the other way round. E.g. "The lecturer speaks too quietly." The same diagram would now have to be interpreted completely differently. Agreement is a poor result.

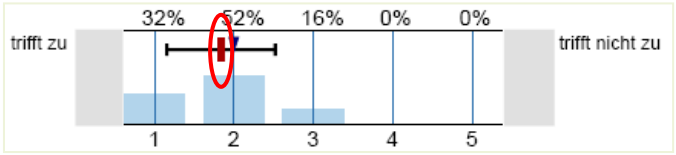
Positive or negative wording of questions

In some questionnaires, questions are formulated partly positively and partly negatively. This is intended to ensure that the questionnaire is filled out consciously. Reversing the polarity of the questions ensures that concentration is maintained when completing the questionnaire. It is also possible to check the consistency of the answers by reversing the polarity. If all the boxes on a questionnaire are ticked in the first "applies" box, it is likely that the questionnaire was not completed seriously.

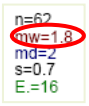
MEAN

The term mean is somewhat imprecise, as it refers to the arithmetic mean. In everyday language, the term average is often used. It is precisely this average that is meant here.

The mean is represented by a vertical red line. In the example graphic, you can see that this is just below the value "2".



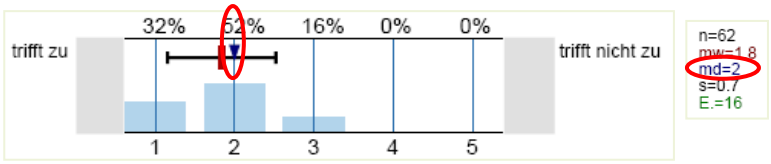
The value is displayed more precisely in the overview of the measurements. It is abbreviated as "**mw**" and shown in the second line. In this example, it is 1.8.



MEDIAN

Less well known than the arithmetic mean (mean value, average) is the median, also known as the central value.

In the example, 20 of the 62 students ticked the first box (rounded to 32%). 32 students ticked the second box (rounded 52%) and the third box was ticked by 16%, i.e. 10 students.



To determine the median, you mentally sort the students in a row. All students stand next to each other. On the left are the 20 students who have ticked box one (numbered from 1-20), to the right are the 32 students who have ticked the second box (numbered from 21-52) and on the far right are the 10 students who have ticked box three (numbered from 53-62). 62 students are now standing next to each other, sorted in ascending order according to the box they have ticked. If you ask the person in the middle of the row which box they ticked, the answer is the median. As there is no middle with 62 people - the middle is between the 31st and 32nd person - you ask both the 31st and 32nd person and calculate the average. As both the 31st and 32nd respondent in the example ticked the second box, the median is therefore "2".

The median is marked with "**md**".

Excursus: Why the median?

The median is more robust against outliers than the mean value (arithmetic mean). It is also the most suitable measure for ordinally scaled data. But what does that mean?

Outliers:

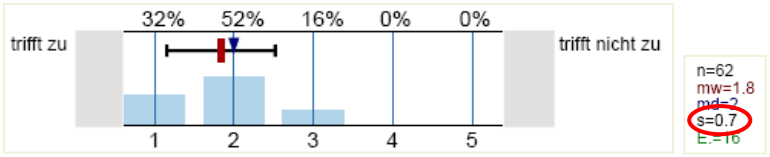
Suppose 20 people are asked about their income and they are all more or less close together. In this case, the mean value is a suitable measure to describe the income of the 20 people. However, if one of the 20 people were replaced by someone with a much higher income, the mean value would be greatly increased by this "outlier". The mean value would no longer represent the sample well. The outlier pulls the mean upwards. The median remains constant despite the outlier. In the case of scale questions, there are no outliers, as there are only a limited number of boxes (usually five). Therefore, the following second argument counts.

Ordinal scaled variable:

The checkbox scales in questionnaires can be understood as ordinally scaled data. This means that they can be put in a clear order, with a scale from "agree" to "disagree", each box further to the right symbolizes less agreement than the box to the left. However, it can be assumed that the mental distance between the first and second or the second and middle box is not the same. It is only a small step from a medium assessment (partly) to a tendency to agree. However, it is a longer way from slight agreement to full agreement. In such cases with only apparently equal intervals (most grades or checkbox scales), the median should be used. Strictly speaking, the mean is not the correct measure because it is based on the assumption of equal distances between the answer options. In practice, however, this fact is almost always ignored. This is why the median is also quite unknown, although it would be "appropriate" in many cases.

STANDARD DEVIATION

One measure of dispersion is the standard deviation. What does dispersion mean? The dispersion indicates how much the answers deviate from the mean value. If all students have ticked the same box, then the dispersion is "0". The results do not scatter, as every student ticked the same box and therefore also the mean value. However, the less the students "agreed", the more the answers scatter, i.e. the greater the standard deviation.

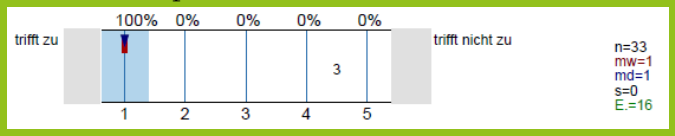


However, a large spread can arise not only due to a different assessment of the content by the students, but also if the question was not asked clearly. Questions with a large spread can be discussed preferentially when giving feedback on the results to the students in order to clarify the reason for the large spread. Is it due to a different assessment of the question or a different understanding of the question that the answers vary so widely?

The standard deviation is abbreviated to **"s"**.

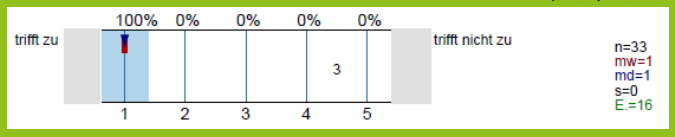
An example of the standard deviation

Suppose 33 students had responded to the statement "The lecturer gives the speakers feedback on their presentation" as follows:



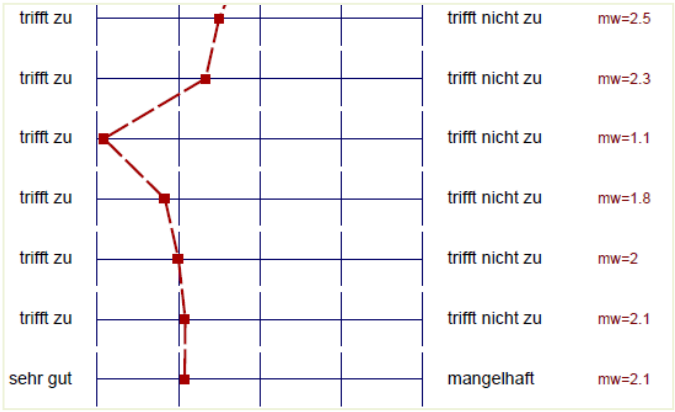
No one deviated from the mean. The standard deviation is therefore s=0. Or to put it another way, on average the students deviate from the mean by zero boxes with their cross.

For the example "The lecturer speaks loudly and clearly", the dispersion is s=0.7. The mean value was given as 1.8. However, as above, not all students answered the same. On average, the students deviated from the mean value by 0.7 boxes for this question.



PROFILE LINE OF ALL SCALE QUESTIONS

The profile line is a special way of displaying mean values. While it is difficult to compare mean values in the bar charts, this is much easier in the profile line.



All scale questions are listed one below the other. The respective mean value is shown in a field to the right. A mean value of 1.0 is on the far left. A mean value of 5.0 is on the far right (with five checkboxes). All other values are divided in between. In this way, outliers can be identified very quickly. In the example graphic, for example, the third question immediately catches the eye.

Caution, trap!

The profile line must be interpreted cautiously for several reasons:

1. trap

Questions can be formulated differently, even if the labels of the poles are the same (e.g. "applies" and "does not apply"). Depending on whether the question is formulated positively or negatively, a mean value of 1.1 - as in the third question in the example profile line - can be an excellent or a very poor value. It is also possible that the mean value represents a positive statement. For example, in the question about the speed of the presentation with the poles "too fast" and "too slow". In this case, the mean value 3 is a good result, while 1 and 5 are poor results.

2. trap

A mean value can always be arrived at in different ways:

A) For example, a mean score of 3.0 can arise because 62 students all ticked exactly the middle box.
B) However, the value 3.0 also arises if 31 students have ticked the first box and 31 students have ticked the fifth box.
C) Or 15 students ticked the first box, 15 students the second, 2 students the third, 15 students the fourth and 15 students the fifth. Here, too, the mean value is 3.0!

The mean value suggests that the students rated the question in the middle range.

In example A, this also applies in full. In example B, the mean value quickly leads to the "wrong track". One half fully agrees with the statement, the other half completely rejects it. Nobody gives a "partly-partly" answer, even if the mean value suggests this.

The following applies: The profile line helps to quickly identify outliers. However, if in doubt, it makes sense to take a look at the bar chart. There it quickly becomes clear how the respective mean value came about.

FURTHER QUESTIONS?

Mirko Birkenkamp

RUHR-UNIVERSITÄT BOCHUM
DEZERNAT 1 – HOCHSCHULENTWICKLUNG UND STRATEGIE
ABTEILUNG 1 – LEHRE, INFORMATIONS- UND QUALITÄTSMANAGEMENT
Tel.: 0234 - 32- 25986
E-Mail: evasys@rub.de
www.rub.de/evasys (Studentische Lehrveranstaltungsbewertung)