CHAPTER 2 – DATA COLLECTION AND PROCESSING (DCP)

<table>
<thead>
<tr>
<th>ASPECT 1</th>
<th>ASPECT 2</th>
<th>ASPECT 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>LEVELS/MARKS</td>
<td>Recording raw data</td>
<td>Processing raw data</td>
</tr>
<tr>
<td>COMPLETE/2</td>
<td>Records appropriate quantitative and associated qualitative raw data, including units and uncertainties where relevant.</td>
<td>Processes the quantitative raw data correctly.</td>
</tr>
<tr>
<td>PARTIAL/1</td>
<td>Records appropriate quantitative and associated qualitative raw data, but with some mistakes or omissions.</td>
<td>Processes quantitative raw data, but with some mistakes and/or omissions.</td>
</tr>
<tr>
<td>NOT AT ALL/0</td>
<td>Does not record any appropriate quantitative raw data or raw data is incomprehensible.</td>
<td>No processing of quantitative raw data is carried out or major mistakes are made in processing.</td>
</tr>
</tbody>
</table>

2.1 RECORDING RAW DATA

'Ideally, students should work on their own when collecting data.

When data collection is carried out in groups, the actual recording and processing of data should be independently undertaken if this criterion is to be assessed.'

© IBO 2007

When being assessed in this aspect, you should not be told by your IB Physics teacher how to record the raw data i.e. you will not be given a pre-formatted table.

RAW DATA

Physics is the most quantitative of all sciences and in many investigations, you will be expected to measure physical quantities. In IB Physics, measurements are considered complete only when the associated uncertainty is stated. Measurements consist of the best estimate and its associated uncertainty is the range of values within which the correct value lies. Therefore all measurements must be expressed as $x \pm \Delta x$ where $x$ is the best estimate of the measurement and $\Delta x$ is the uncertainty. For example, the time for five complete oscillations of a simple pendulum can be expressed as $(5.3 \pm 0.2)$ seconds.

© S.T.A.R. 2007 This page may only be legally copied under the conditions of sale to the purchasing school.
Uncertainties are associated with all raw data and an attempt should always be made by you to quantify uncertainties. For example, when there is an uncertainty in a time measurement because of reaction time, you must estimate the magnitude of the uncertainty.

PRESENTING RAW DATA

Most of the observations and measurements that you will record need to be presented in a table of results. It is important that your table of results is effective and easy to read. Therefore, when designing your results table, it should be able to clearly communicate the relationship between the dependent and independent variables.

Here are a few guidelines that may help you design a clear and efficient table of results.

- The independent variable column should be left most, the dependent variable column next.
- Often there is a need to do repeat readings. Sub-divide the column for the dependent variable to reflect the number of trials conducted.
- Column headings: The heading of each column must clearly describe the data the column represents, the units in which the measurement is made and an indication of the uncertainty.
- Where appropriate, record and show the value of any controlled variable.
- Where appropriate, indicate the source of uncertainty in the measurement.

UNCERTAINTY AND PRECISION

- Readings made directly from measuring instruments should be given to the number of decimal places that is appropriate for the measuring instrument used, that is, the number of significant digits must clearly reflect the precision of the measuring instrument.
- Be consistent with the degree of precision for each column of raw data. (i.e. if data is precise up to 1 decimal place, record 2.0, not 2).
<table>
<thead>
<tr>
<th>MEASURING INSTRUMENT</th>
<th>SMALLEST DIVISION</th>
<th>PROBABLE UNCERTAINTY</th>
<th>EXAMPLES OF THREE CONSECUTIVE READINGS</th>
</tr>
</thead>
<tbody>
<tr>
<td>METRE-RULE</td>
<td>0.1 cm</td>
<td>± 0.05 cm</td>
<td>10.20 cm, 10.25 cm, 10.30 cm</td>
</tr>
<tr>
<td>VERNIER CALIPER</td>
<td>0.01 cm</td>
<td>± 0.01 cm</td>
<td>6.23 cm, 6.24 cm, 6.25 cm</td>
</tr>
<tr>
<td>MICROMETER SCREW GAUGE</td>
<td>0.01 mm</td>
<td>± 0.01 mm</td>
<td>5.15 mm, 5.16 mm, 5.17 mm</td>
</tr>
<tr>
<td>ELECTRONIC BALANCE</td>
<td>0.01 g</td>
<td>± 0.01 g</td>
<td>179.99 g, 180.00 g, 180.01 g</td>
</tr>
<tr>
<td>ANALOGUE STOPWATCH</td>
<td>0.1 s</td>
<td>± 0.05 s</td>
<td>2.00 s, 2.05 s, 2.10 s</td>
</tr>
<tr>
<td>DIGITAL STOPWATCH</td>
<td>0.01 s</td>
<td>± 0.01 s</td>
<td>15.19 s, 15.20 s, 15.21 s</td>
</tr>
<tr>
<td>LIQUID-IN-GLASS THERMOMETER</td>
<td>1 °C</td>
<td>± 0.5 °C</td>
<td>30.0 °C, 30.5 °C, 31.0 °C</td>
</tr>
<tr>
<td>DIGITAL THERMOMETER</td>
<td>0.01 °C</td>
<td>± 0.01 °C</td>
<td>29.89 °C, 29.90 °C, 29.91 °C</td>
</tr>
<tr>
<td>DIGITAL AMMETER</td>
<td>0.01 A</td>
<td>± 0.01 A</td>
<td>1.49 A, 1.50 A, 1.51 A</td>
</tr>
<tr>
<td>ANALOGUE AMMETER (0 – 3 A)</td>
<td>0.1 A</td>
<td>± 0.05 A</td>
<td>2.60 A, 2.65 A, 2.70 A</td>
</tr>
</tbody>
</table>

_Figure 201  Common instruments and their uncertainty_

Note again that the measurement and the uncertainty must be expressed to the same degree of precision, that is, the same number of decimal places. See Figure 201.
EXAMPLE: DATA COLLECTION

**Independent variable is the first column**

**Length of the pendulum / cm**

<table>
<thead>
<tr>
<th>Length (± 0.05 cm)</th>
<th>Time to complete 5 oscillations / s (± 0.21 s)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Trial 1</td>
</tr>
<tr>
<td>40.00</td>
<td>6.54</td>
</tr>
<tr>
<td>50.00</td>
<td>7.47</td>
</tr>
<tr>
<td>60.00</td>
<td>7.87</td>
</tr>
<tr>
<td>70.00</td>
<td>8.93</td>
</tr>
<tr>
<td>80.00</td>
<td>8.97</td>
</tr>
<tr>
<td>90.00</td>
<td>9.63</td>
</tr>
<tr>
<td>100.00</td>
<td>9.97</td>
</tr>
</tbody>
</table>

**Clear headings with units and uncertainties**

**Precision of uncertainty and measurement is consistent e.g. two decimal places**

**Precision of measurements is consistent.**

**Column subdivided to show repeat readings**

**Controlled variables**

- Mass of the pendulum bob = 300 g
- Angle of release = 5°

**Error explained**

- Uncertainty in time is due to reaction time (± 0.20 s) and instrument limitation (±0.01 s)

**Figure 202 Sample data for the investigation “Period of a pendulum”**

**Good practices in data collection**

- It is important that you keep an accurate record of what you do in the laboratory. What you use is a matter of personal choice but it is suggested that you keep a laboratory notebook for this purpose.

- It is important that you have drawn the results table before the experiment is performed. It is not good practice to initially record your results on a piece of rough paper and transfer it later to an organized data table. This runs the risk of copying errors or worse, of losing the paper.

- Carry out your measurements carefully and as accurately as possible. No amount of careful calculation after the investigation is over can compensate for improperly recorded data. If the data are incorrect, the aim of the investigation is completely lost.

- When collecting data, it is good routine to do a practice run before doing the actual measurement. This will give you a feel of the measurements to be recorded and of the possible problems that may be encountered. This will also help you decide whether you need to do repeat readings.
2.2 **PROCESSING RAW DATA**

“Data processing involves, for example, combining and manipulating raw data to determine the value of a physical quantity (such as adding, subtracting, squaring, dividing), and taking the average of several measurements and transforming data into a form suitable for graphical representation. It might be that the data is already in a form suitable for graphical presentation, for example, light absorbance readings plotted against time readings. If the raw data is represented in this way and a best-fit line graph is drawn and the gradient determined, then the raw data has been processed. Plotting raw data (without a graph line) does not constitute processing data.”

“Most processed data will result in the drawing of a graph showing the relationship between the independent and dependent variables.”

When being assessed in this aspect, you should not be told by your IB Physics teacher:

- how to process the data
- what quantities to graph/plot.

Data processing in the IB Physics course can take a variety of forms. Listed below are ways by which you can process data. Note that you may not need to do all of these methods with every investigation that you will be doing.

**CALCULATING THE AVERAGE**

When measurements are repeated, you need to calculate the average of the values. The calculated average value will then be used as the value of the measurement in the subsequent stages of the data processing. Also, the average value cannot be more precise than the measurements, thus it should be rounded off to the same number of decimal places as the measurements.

The method for calculating the uncertainty in repeated measurements is described in Section 8.1.7.

**CALCULATING A PHYSICAL QUANTITY**

In some investigations, you will need to directly calculate some physical quantity by using measured variables. At times, this physical quantity can be derived from the gradient of the graph. As an example, you might be asked to investigate a factor which affects the electrical resistance of a conducting wire. If you have a graph of electrical resistance against length you would be able to calculate the resistivity of the wire from the gradient.

What is important is that you show the working clearly and you round off your final value to the correct number of significant figures. You would also need to show how the uncertainties are propagated in the calculation. For this, it is most helpful if you change the absolute uncertainties into percentage uncertainties.
Calculation involving measurements and uncertainties entail some rules that you need to remember. These rules are described in Sections 8.1.6 – 8.1.8 which cover uncertainties and the propagation of errors.

**PLOTTING GRAPHS**

Graphing is an essential part of data processing in the IB Physics course. Plotting a graph is a way of:

- averaging the data so that the effects of random errors are minimized.
- identifying anomalous points.
- eliminating some systematic errors.

In many cases, you would need to plot a graph to determine or verify the relationship between the independent and dependent variable, and the graph is usually (although not always) a straight line. In most graphs, you must always plot the best-fit straight line or curve. The line of best fit and how to draw it are described in Section 8.2.1.

The graph that is easiest to analyse and interpret is that which has a straight line. However, when the values of the dependent and independent variables are plotted, not all trends will produce a straight line. The analysis of a graph with a curved line is not as simple as that of a straight line graph. Fortunately, it is possible, by carefully selecting the variables to be plotted, to 'linearise' a curved line graph.

If the equation relating the variables is known, you must be able to select the appropriate variables to plot to obtain a straight line. When the equation relating the two variables is not known, you may use a log-log graph to obtain a straight line graph and use its gradient to deduce the relationship. The method of linearising a graph is described fully in Section 8.2.5.

**CALCULATING AND INTERPRETING THE GRADIENT**

The determination of the gradient of the graph is an essential component of data processing. Note that if a graph is present, data processing is not considered complete unless the gradient is obtained. The method of calculating the gradient is shown in Section 8.2.2. In many cases, you may also need to interpret the significance of the gradient or use it to obtain another physical quantity. Examples of interpreting the significance of a gradient are shown in Section 8.2.4.

Note that the full procedure of plotting a graph is discussed in Section 8.2 – Graphs.
2.3 Presenting Processed Data

The processing of data needs to be presented in a clear and efficient manner. All key stages must be presented sequentially so that the pathway to the conclusion is clear to the reader. Also, tabulate the processed data if possible.

A good data processing presentation will include:

- display of key calculations in a clear and sequential manner.
- the propagation of errors in the calculated values (Section 8.1.8).
- graph with the appropriate scales, clearly labeled axes with units (Section 8.2.1).
- uncertainty bars in straight line graphs (Section 8.2.1).
- lines of minimum and maximum gradients (Section 8.2.3).
- determining the uncertainty of the best straight line gradient. (Section 8.2.3)

**EXAMPLE: DATA PROCESSING FOR THE INVESTIGATION “PERIOD OF A SIMPLE PENDULUM”**

Rationale for the data processing.

The relationship between the length $l$ and the period $T$ of the pendulum is given by the equation: $T = 2\pi \sqrt{\frac{l}{g}}$

To obtain a straight-line graph, one needs to plot $T^2$ against length $l$.

\[
T^2 = 4\pi^2 \frac{1}{g}
\]

\[
T^2 = k \frac{1}{g} \quad \text{where} \quad k = 4\pi^2
\]
Length is converted to the necessary unit. Note that the uncertainty and the data have the same precision.

<table>
<thead>
<tr>
<th>Length of the pendulum / m ± 0.005 m</th>
<th>Average time for five oscillations / s ± 0.21 s</th>
<th>Period / s ± 0.03</th>
<th>Period $^2$ / s² ± 0.1</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.40 00</td>
<td>6.66</td>
<td>1.33</td>
<td>1.8 ± 0.1</td>
</tr>
<tr>
<td>0.50 00</td>
<td>7.07</td>
<td>1.41</td>
<td>2.0 ± 0.1</td>
</tr>
<tr>
<td>0.60 00</td>
<td>7.86</td>
<td>1.57</td>
<td>2.5 ± 0.1</td>
</tr>
<tr>
<td>0.70 00</td>
<td>8.64</td>
<td>1.73</td>
<td>3.0 ± 0.2</td>
</tr>
<tr>
<td>0.80 00</td>
<td>8.95</td>
<td>1.79</td>
<td>3.2 ± 0.2</td>
</tr>
<tr>
<td>0.90 00</td>
<td>9.57</td>
<td>1.91</td>
<td>3.6 ± 0.2</td>
</tr>
<tr>
<td>1.0000</td>
<td>10.10</td>
<td>2.02</td>
<td>4.1 ± 0.2</td>
</tr>
</tbody>
</table>

Uncertainty rounded off to one significant figure. The data value is also rounded off so that it has the same precision as the uncertainty.

Calculated values have the same precision as the measured values.

Figure 203  An example of a table of processed results.
Gradient = 3.9 ± 0.6 s² m⁻¹

Uncertainty is obtained by calculating the average deviation between the gradient of the line of best fit and the maximum or minimum gradient.

Figure 204  Graph of Period² against length
EXAMPLE: CALCULATION FOR THE GRAVITATIONAL ACCELERATION CONSTANT $g$ FROM THE GRADIENT

From the graph, the gradient is equal to $\frac{\Delta T^2}{\Delta l}$.

Since $T^2 = 4\pi^2 \frac{l}{g}$, then by comparing it to the gradient, $g$ is equal to

$$g = \frac{4 \pi^2}{\text{gradient}} = \frac{4 \pi^2}{3.9} = 10.1$$

To obtain the uncertainty in $g$, use the above formula with the maximum and minimum values of $g$.

Maximum value: $$g = \frac{4 \pi^2}{\text{gradient}} = \frac{4 \pi^2}{3.2} = 12.3$$

Minimum $g = \frac{4 \pi^2}{\text{gradient}} = \frac{4 \pi^2}{4.6} = 8.6$

Uncertainty in $g = \frac{12.3 - 8.6}{2} = 1.9$

Therefore $g = (10.1 \pm 1.9) \text{ m/s}^2$

or more realistically, the result should be quoted as

$g = (10 \pm 2) \text{ m/s}^2$. 
<table>
<thead>
<tr>
<th><strong>ASSESSMENT CRITERIA</strong></th>
<th><strong>EVIDENCE REQUIRED</strong></th>
<th><strong>WHAT YOU MUST DO</strong></th>
</tr>
</thead>
</table>
| Records appropriate quantitative and associated qualitative raw data, including units and uncertainties where relevant. | Clear, organised and accurate presentation of results and observations. | Neat, orderly table, boxed in with  
• clear, descriptive and concise headings with suitable units  
• logical sequence of the readings with the appropriate number of significant figures, reflective of the degree of precision of the apparatus. |
| Identification of significant sources of error. | State what the uncertainties are and estimate the uncertainties. Ensure that the readings and the uncertainty have the same number of decimal places. |
| Processes the quantitative raw data correctly. | Appropriate processing of the data. | Show that the calculations performed have a scientific basis and have links to the aim(s) of the investigation.  
The level of precision of the processed data should be consistent with that of the raw data. |
| Where applicable, a suitable graph. | In most cases, the “suitable graph” is a straight line graph. Therefore, the quantities plotted are such that they would produce a straight line.  
When the gradient is calculated, a large “triangle” is used and shown in the graph.  
State how the gradient is related to the formula. |
| Presents processed data appropriately and, where relevant, includes errors and uncertainties. | Clear, organised presentation of the processing of data. | Show clearly all necessary calculation in a logical sequence.  
Show sample calculations if the process is repeated.  
Where appropriate, present a neat and organised table of processed data with headings and units and uncertainties.  
Ensure that the graph has appropriate scales, labelled axes with units, and accurately plotted data points marked by crosses, with a suitable best-fit line or curve.  
Use a sharp pencil when plotting points and drawing the best-fit line or curve. |
| Clear, organised presentation of the processing of associated errors. | Show how the uncertainties are propagated in the calculations.  
If a straight-line graph is present;  
• draw the uncertainty bars for the first and last data points.  
• explain where the uncertainties are not significant.  
• draw lines of maximum and minimum gradients – use dashed lines instead of solid lines.  
• calculate the uncertainty in the best straight-line gradient. |

*Figure 205  Summary of Data Collection and Processing criterion*