Table 2 (extended). Category framework used for the item analysis of this article. For categories that are less obvious in definition, examples from IBO exam items are presented (accessible at: https://www.ibo-info.org/en/info/papers.html). Some categories (i.e., the readability formulas, taxonomic order) did not require a rating, as their outcome is clearly defined and does not depend on the interpretation of a rater. “/” = not applicable.

| Category                     | Sub-categories | Traits | Example for rating instructions (Note: Full coding instructions include different inclusion- and exclusion criteria, as well as examples.) | Example items or item features | Authors                  |
|------------------------------|----------------|--------|-------------------------------------------------------------------------------|-----------------------------|---------------------------|
| Areas 1: Formal item characteristics |                |        |                                                                               |                             |                           |
| Response type                |                |        |                                                                               |                             |                           |
| Single choice                |                |        |                                                                               |                             |                           |
| Multiple choice              |                |        |                                                                               |                             |                           |
| Labeling                     |                |        |                                                                               |                             |                           |
| Multiple true / false (MTF)  |                |        | e.g., MTF: Code if several answering options have to be identified as right or wrong. |                             |                           |
| Fill in blanks               |                |        |                                                                               |                             |                           |
| Matching                     |                |        |                                                                               |                             |                           |
| Constr. response             |                |        |                                                                               |                             |                           |
| Practical tasks              |                |        |                                                                               |                             |                           |
| e.g., MTF: The student has to indicate if each of four answering options is true / false. | |                           |
| Marso and Pigge (1991)       |                |        |                                                                               |                             |                           |
| Language:                    | Flesch-Kincaid |        | e.g., 12.3 (US grade level) (computerized calculation)                         | The item's text has a reading difficulty suitable for 12th grade | Kincaid et al. (1975)    |
| Readability                  |                |        |                                                                               |                             |                           |
## Area 2: Contents & practices

| Disciplinary core ideas | Structure & function | Code if the item concerns a relation between structure and function. | The item asks students to analyze reasons for enzyme-mediated reaction velocity in the presence / absence of different enzyme inhibitors. |
|-------------------------|----------------------|---------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------|
|                         | Structure & function | Code if the item concerns regulatory processes.                           | IBO 2013, Theory 1, Task 6
The item asks students to determine the regulatory effects of different antibiotic treatments over time. |
|                         | Steering & regulation |                                                                          | IBO 2013, Theory 1, Task 16
KMK (2004); NGSS (2013)
The item asks students to analyze leaf temperatures at different water supplies with regards to energy transfers to and from the leaf. |
|                         | System               | Code if the item requires the analysis of matter transformation / transfer or energy transfer. | IBO 2013, Theory 1, Task 28
The item asks to analyze electric organ discharges used by different fish species as a means of communication. |
|                         | Transfer & transformation of matter & energy | Code if the item concerns taking up, creating, adapting or exchanging information between systems or system compartments. | IBO 2013, Theory 1, Task 45
The item asks to analyze electric organ discharges used by different fish species as a means of communication. |
|                         | Information & communication |                                                                          | IBO 2013, Theory 1, Task 45
The item asks to analyze electric organ discharges used by different fish species as a means of communication. |
| Disciplinary core ideas (continued) | Reproduction | Development | Variability & adaptation | Phylogenesis & relatedness | Nature of Science (NOS) |
|-----------------------------------|--------------|-------------|---------------------------|---------------------------|------------------------|
| Code if the item is related to description of (a-) sexual reproduction. | | Code if the item concerns biological variability and the reasons for this variability and/or aspects of adaption. | | Code if the item concerns similarities and variability of life forms as a results of their phylogenetic development. | Code if the item concerns at least one of the following: (a) intents or processes of scientific activity, (b) ideas that help explain nature or produce products have been derived (c) applications of science with regards to ethical / environmental implications (d) quality of different information sources, (e) certainty of scientific knowledge, its justification or its development over time. |
| The item asks to derive allele frequency fluctuations in a population with specified reproduction parameters. | | | | | The item asks students to compare different theoretical explanations for a phenomenon and explain the strengths of each one. |
| IBO 2013, Theory 2, Task 32 | | | | | Conley et al. (2004); Harlen et al. (2013, 2015) |
| Students are asked to describe the relationship between hemoglobin O₂-affinity in species of different habitats and environmental O₂ pressures. | | | | | |
| KMK (2004); NGSS (2013) | | | | | |
| IBO 2013, Theory 1, Task 19 | | | | | |
| Students are asked to draw conclusions from different equations that represent evolutionary principles | | | | | |
| IBO 2017, Theory 1, Task “Evolutionary genetics” | | | | | |
| Scientific practices                  | Code if the item requires students to formulate / select suitable research question, hypotheses or problems. | Students evaluate the correctness of different hypotheses for a given experiment. |
|--------------------------------------|---------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------|
| Asking questions & defining problems | Code if the item requires students to design, build, use, critique or revise a model or to select options that reflect these processes. | IBO 2009, Theory B, Task 15                                                                 |
| Developing & using models            | Code if the item requires students to plan or carry out investigations or to selection options that reflect these processes. | Students are required to use a model of double strand brakes to predict allele frequencies |
| Planning & carrying out investigations| Code if the item requires students to analyze data (either raw, in graph form, or in summarizing representations such as flow charts). | IBO 2013, Theory A, Task 31                                                      |
| Analyzing & using data               | Code if the item asks students to design / run / evaluate simulations in a quantifying manner or to analyze / conduct / evaluate statistics – either open-ended or by selecting options that reflect these processes. | Students analyze the effectiveness of different platelet aggregation inhibitors over time from line diagrams. |
| Using mathematics & computational thinking |                                                                                                           | IBO 2013, Theory A, Task 18                                                      |
|                                      |                                                                                                           | Students calculate means and variances across different replications in experimental data they collected and conduct t-tests on their data. |
|                                      |                                                                                                           | IBO 2013, Practical 3 “Evolutionary Ethology”                                     |

**IBO 2013, Practical 3 “Evolutionary Ethology”**
| Scientific practices (continued) | Constructing explanations & designing solutions | Engaging in argumentation from evidence | Obtaining, evaluating and communicating information |
|----------------------------------|-----------------------------------------------|----------------------------------------|--------------------------------------------------|
|                                 | Code if students construct explanations of phenomena based on claim, evidence and reasoning or if they select options that reflect these processes. | Code if the item asks to use appropriate and sufficient evidence (out of several options) and use scientific reasoning (e.g., weighing of data quality, using scientific principles) to support of oppose claims, explanations or design solutions. | Code if students compare or integrate different data or information sources or if they assess the information’s quality regarding its reliability or validity. |
|                                 | Students use evidence from Western blotting to provide explanations for the role of different enzymes in Salmonella growth.  
*IBO 2017, Theory 1, Task “Acetylation”* | The student compares different data sources regarding an overarching question/problem on global warming, weighing them to derive a solution.  
*NGSS (2013); KMK (2004)* | Students evaluate data and information for different DNA sequencing methods regarding their validity in different settings.  
*IBO 2017, Theory 1, Task “Sequencing”* |

| Context authenticity | Authentic pieces of life science research | Reference to students’ life world |
|----------------------|------------------------------------------|----------------------------------|
| Yes / No             | Code if the item engages student in excerpts, findings or summaries from pieces of scientific research in the life sciences. | Students match different floral patterns and pollinator insect specializations.  
*IBO 2013, Practical 2 “Plant Physiology, Morphology and Ecology”* |
|                      | Students are presented with findings by Dame Leyer. They analyze auxin oxidation rates from extracts of tumorous galls to derive implications of this research.  
*IBO 2017, Theory 1, Task “Oak galls”* | Weiss and Müller (2015) |
| Biological domain | e.g., Ethology: Code if the item concerns aspects of behavior. |
|-------------------|-------------------------------------------------------------|
|                    | e.g., Ethology: Students analyze data from an experiment on human wake-up times in an environment without external time cues. |
|                    | IBO 2013, Theory B, Task 40 |

| Taxonomic order | e.g., Primates / (internet search) |
|-----------------|-----------------------------------|
|                 | E.g., An item concerns human biology. |
|                 | Various |

| Organizational level | e.g., Biome: Code if the item has reference to larger geographic areas like continents. |
|----------------------|------------------------------------------------------------------|
|                      | E.g., Organ: Students compare digestive systems of different invertebrates. |
|                      | IBO 2013, Theory B, Task 43 |
|                      | Solomon, Berg, and Martin (2011) |
### Area 3: Cognitive aspects

| Types of knowledge | Factual | Conceptual | Procedural | Metacognitive |
|--------------------|---------|------------|------------|--------------|
|                    | Facts   | Classifications & generalizations | Specific skills & algorithms | Strategic thinking |
|                    | Terms   | Theories, Models & Structures | Techniques & methods | Knowledge about cognitive tasks, contextual and conditional knowledge |
|                    | Specifics |                         | Knowing when to use a specific procedure | Self-knowledge |

#### Factual
- **Code if the item requires students to exhibit basic knowledge element of biology, i.e., facts / terms (e.g., technical terms).**
  - Students indicate, which phase of the cell cycle includes DNA replication.
  - *IBO 1993, Theory exam, Task 24*

#### Conceptual
- **Code if the item requires students to address the interrelation between basic elements of biology. This includes classifications (e.g., climate periods), principles (e.g., probability) or theories/model/structures (e.g., theory of evolution).**
  - Students analyze historical changes in abundance of three specialized midge species. To solve the item, they interpret the data in the light of past climatic and ecological conditions.
  - *IBO 2013, Theory A, Task 43*

#### Procedural
- **Code if the item requires students to know how or when to do a certain procedure. This can include biological skills (e.g., titration) or general techniques (e.g., how to formulate suitable hypotheses) or students can be required to establish criteria for which procedure is most suitable.**
  - To determine the density of different Trypanosome strains, students create dilutions, use pipettes to transfer solutions and buffers and calculate means and standard deviations.
  - *IBO 2013, Practical 1 “Molecular Cell Biology”*

#### Metacognitive
- **Code if the item requires students to demonstrate strategic thinking, knowledge about cognitive demands of specific tasks or if the student has to exhibit self-awareness (e.g., about her/his own abilities).**
  - Students receive points for stopping at an intermediate point of solving a longer problem to reflect what they have learned, how they should continue or what methods it could be valuable to learn before continuing.
  - *No clear examples amongst the analyzed IBO items.*

*Anderson and Krathwohl (2001); Bloom et al. (1956)*
| Cognitive processes | Remember | Recognize | Recall |
|---------------------|----------|-----------|--------|
| Understand          |          | Interpret | Exemplify |
|                     |          | Classify  | Summarize |
|                     |          | Infer     | Compare  |
|                     |          | Explain   |          |

- **Remember**: Students indicate, which phase of the cell cycle includes DNA replication.  
  *IBO 1993, Theory exam, Task 24*

- **Understand**: Students have to compare their knowledge of cell organelles in different cell types.  
  *Anderson and Krathwohl (2001)*;
  *Bloom et al. (1956)*

- **Apply**: To determine heat exchange with the environment in differently sized animals, students apply geometric formulas for ratios between length, surface and volume.  
  *IBO 2017, Theory B, Task “Organism Scaling”*

- **Execute**: Code if the student has to recognize individual pieces of knowledge (e.g., structure of an amino acid formula) or recall them (e.g., provide the sum formula for ribose), but not any of the higher-order cognitive processes.

  Code if the item requires students to construct meaning by interpreting (e.g., by rephrasing), exemplifying (e.g., providing an example), classifying (e.g., by categorizing), summarizing (e.g., by finding the main point in a text), inferring (e.g., by extrapolating an observed relationship between variables), comparing (e.g., two analogous structures), or explaining (e.g., by establishing cause-effect-relationships).

  Higher-order cognitive processes must not be realized by the same item.

- **Use**: Code if the item requires students to apply a procedure (e.g., calculating variance) to a familiar or unfamiliar task.
### Cognitive processes (continued)

| Analyze | Differentiate | Organize | Attribute |
|---------|--------------|---------|-----------|
|         |              |         |           |

**Analyze**

- Differentiate
- Organize
- Attribute

Code if the item requires students to analyze parts of a content / system and determine their relationships, either by differentiating (e.g., identifying trends in data), by organizing (e.g., setting up a phylogenetic tree) or by attributing (e.g., behavioral traits to an animal).

Students analyze the effectiveness of different platelet aggregation inhibitors over time from line diagrams.

*IBO 2013, Theory A, Task 18*

**Evaluate**

- Check
- Critique

Code if the item requires students to check biological contents / procedures / data for internal consistency (e.g., regarding conclusions that can be drawn from a data set), external consistency (e.g., implication of findings in the context of a societal problem or laws), or effectiveness (e.g., reliability / validity of applied procedures).

Students evaluate data from two different diagnostic approaches for bacterial infections in artificial joints with focus on the tests’ respective reliability (here: false-positives) and suitability as a large-scale diagnostic procedure.

*IBO 2013, Theory B, Task 13*

**Create**

- Generate
- Plan
- Produce

Code if the item requires students to re-organize elements (knowledge, physical- or abstract objects) into a coherent or functional new unit. Students can do so by generating (e.g., hypotheses, questions), designing (e.g., research plans, model sketches) or constructing (e.g., preparing a research set-up, constructing / improving a model).

Based on prior morphological data they analyzed for a number of mammal species, students design a phylogenetic tree for these species.

*IBO 2013, Practical 2*  
“Comparative and Functional Biosystematics”
## Area 4: Use of representations

| Representation: type | Description | Example |
|---------------------|-------------|---------|
| Descriptive         | Code if the items presents text only (no graphics or formulas). | An item with text only. |
| Depictive-logical   | Code if the item presents quantifying relationship in visual form. | An item encompassing a table, map, graph, or similar object. (Schnottz 2005) |
| Depictive-realistic | Code if the item presents visual representation oriented after reality. | An item with a drawing, photo, simulation or similar object. |
| Symbolic            | Code if the item presents formula. | An item with a formula. |

| Representations: functions | Description | Example |
|----------------------------|-------------|---------|
| Low                        | Code if item presents isolated unit, no processes, no interactions. | E.g., “What is the definition of a gene?” (IBO 1993, Theory, Task 3) |
| Medium                     | Code if some parts of the system are shown, yet no processes indicated. | Students analyze a phylogenetic tree from a methodological perspective - no focus on the underlying temporal scale (Slough and McTigue 2013) |
| High                       | Code if multiple parts of system are shown, helping to build mental model of the system; processes are indicated as change over time. | Students analyze a detailed signaling pathway system for ethanol detection in Drosophila (IBO 2017, Theory 1, Task “Drunk flies”) |
| Captions | Low | Code if neither captions, nor other introductions to the representation are provided. | A graphic without reference or title. |
| --- | --- | | |
| | Medium | Code if a short, summary-like statement is provided in the item stem and/or directly above the graphic. | E.g., *IBO 2017, Theory 1, Task “Shoot branching”* |
| | High | Code if a detailed graphic caption (either above the figure or in the item stem) is provided that names relevant elements (e.g., axes, elements, symbols). | The item stem addresses both the following figure and provides detailed information about its elements. |
| | | | *IBO 2013, Theory B, Task 5* |
| Representations: functions (continued) | Decorative | Code if the graphic is independent from the relevant information used to solve the item, i.e., it doesn’t support the meaning of the text. | The item speaks about findings from a researcher, for whom a portrait picture is added. |
| | Semantic relationship: text & graphic | Representational | Code if the graphic shows what is written in the text, but without providing additional information. | *IBO 2017, Theory 1, Task “Epigenetics of flowering”* |
| | Organizational | Code if the graphic adds coherence to the item by putting the information in a different frame (e.g., greater picture). | A drawing visualizes the experimental setup for a light exposure experiment described in the accompanying item stem. |
| | | | *IBO 1993, Theory, Task 31* |
| | | | An item about the haem catabolism. Drawings of the molecules involved in the process are shown besides the item stem, but they are not needed for the solution. |
| | | | *IBO 2013, Theory A, Task 12* |

Slough and McTigue (2013)
An item stem introduces zebra fish and the animals’ varying stripes, as well as the physiology behind these markings. The following representation summarizes experimental procedures and findings used to determine the formation of cells that form the stripes.

IBO 2017, Theory 1, Task “Turing patterns”
Assessing High Performers in the Life Sciences: Characteristics of Exams used at the International Biology Olympiad (IBO) and Their Implications for Life Science Education

Sebastian Opitz & Ute Harms

Supplemental Material 2 – Extended Figures
Figure 2 – extended. Shares (%) of items incorporating different of core ideas (KMK, 2004; NGSS, 2013) across six IBO assessment cohorts. Different letters above the columns refer to statistically significant differences between IBO years. They are only displayed for categories were trends across multiple IBO years are discernable.
**Figure 3 - extended.** Shares (%) of items incorporating different of scientific practices (NGSS, 2013) across six IBO assessment cohorts. Different letters above the columns refer to statistically significant differences between IBO years. They are only displayed for categories were trends across multiple IBO years are discernable.