Resources for training trainers

Introducing the
Mastery Rubric for Bioinformatics

PROFESSIONAL GUIDE

Version: 30 June 2022

Introducing the
The Mastery Rubric for Bioinformatics

Overview

This Professional Guide introduces the Mastery Rubric for Bioinformatics (MR-Bi), describing its structure and how it maps performance as learners traverse a developmental trajectory from lower- to higher-order critical-thinking skills. The focus here is on understanding the MR-Bi's key elements, and their potential to inform the assessment of learner development and course design.

Teaching Goals & Learning Outcomes

This Guide outlines the principal components of the MR-Bi. On carefully reading this Guide, and engaging with the reflections and exercises, you will be able to:

- *describe* the general structure of the MR-Bi;
- *list* some of the key Knowledge, Skills and Abilities (KSAs) the MR-Bi was designed to help deliver;
- *define* the developmental stages of the MR-Bi;
- *identify* the Bloom's-level(s) of cognitive complexity, and broad academic level(s), compatible with each stage;
- *describe* how performance and critical thinking change as learners traverse the developmental trajectory;
- *identify* the KSAs that would need to be modified to adapt the MR-Bi to a closely related scientific discipline(s); and
- *pinpoint* your own stage of development in some of the requisite KSAs.

1 Introduction

In the last 20 years, life-science education programmes have had to adapt to reflect the increasingly data-intensive nature of the discipline. Yet, data management, data analytics, scripting, and so on are taught relatively rarely in life-science degree programmes, creating a gap between theory and practice, and fuelling demand for bioinformatics training across all educational levels and career roles[1-6]. This has led to the augmentation of some established **curricula** with short bioinformatics courses, and/or to the development of entire bioinformatics degree programmes.

Bioinformatics education and training requires purposeful integration of discipline-specific perspectives and fundamental knowledge (from computational and life sciences), often in limited time-frames[6,7]. Integrating computational skills and analytical thinking into such courses in systematic and formal ways can therefore be difficult.

To help design education programmes, curriculum guidelines and core bioinformatics **competencies** have been created[8-13]. Missing from such approaches, however, is often the route (developmental trajectory[14]) and time-frame for achieving the competencies; their use in curriculum development has therefore proved challenging[15-18].

To address this issue, a new curriculum-design tool was created: the Mastery Rubric for Bioinformatics (MR-Bi)[19]. Unlike conventional **rubrics**, Mastery Rubrics aim to support the development of specific Knowledge, Skills and Abilities (KSAs) along stages in a developmental trajectory (from uninitiated student to independent practitioner) by describing the performance or behaviours typical of learners at each stage. They hence span the full curriculum rather than individual student assignments[14,19].

The MR-Bi is a framework that supports bioinformatics curriculum and course design, and self-directed learning. It prioritises the development of independence and scientific reasoning, and is structured to allow individuals (regardless of career stage, disciplinary background, or skill level) to locate themselves within the framework.

Based on *The Mastery Rubric for Bioinformatics: a tool to support design and evaluation of career-spanning education and training*[19], this Guide introduces the MR-Bi and forms part of the GOBLET-ELIXIR train-the-trainer resources. Its companion Guide, *Using the Mastery Rubric for Bioinformatics – a Professional Guide*[20] offers insights into how to use the tool in practical education and training scenarios.

2 About this Guide

This Guide provides an overview of the principal features of the MR-Bi. Exercises and Reflections are provided to help readers to consider how the MR-Bi can be used to gauge learners' (and indeed their own) levels of performance and to highlight their requisite training needs, and how this knowledge may be used to support their teaching practice and/or course development. Throughout the text, key terms – rendered in **bold** type – are defined in boxes. Additional information is provided in supplementary boxes and figures.

KEY TERMS

Competencies: multi-dimensional, complex, task-specific behaviours that represent what individuals can do when they bring their knowledge, skills & abilities together appropriately, at the right level(s) for the right application, to achieve a given task

Curriculum: the inventory of tasks involving the design, organisation & planning of an education or training enterprise, including specification of **Learning Outcomes (LOs)**, content, materials & assessments, & arrangements for training teachers & trainers

Learning Outcome (LO): the KSAs that learners should be able to demonstrate after instruction, the tangible evidence that the teaching goals have been achieved; LOs are *learner-centric*

Rubric: in education, a tool used to evaluate & grade student work; often presented in tabular form, rubrics generally contain evaluative criteria, qualitative performance descriptions for those criteria at specific achievement levels & an associated scoring system

3 What is the MR-Bi?

Like other rubrics used in education, the MR-Bi is essentially just a table – in fact, it's a very large 12 by 5 matrix! Just as a conventional rubric encapsulates evaluative criteria to help instructors assess student performance at defined achievement levels, so the MR-Bi also contains descriptions of learner performance; here, however, performance is characterised across career-spanning stages of the full academic spectrum rather than focusing on a particular piece of work. This has ramifications for individuals engaging in professional development (whether to augment existing skills or to acquire new ones), for supervisors aiming to upskill their students, and for instructors developing bioinformatics courses or programmes.

Comparison of the MR-Bi with conventional rubrics

Rubrics are routinely used in a range of educational settings. They're generally used as scoring guides to facilitate rigorous and consistent evaluation of learner performance on a given piece of work. Here, we compare and contrast conventional rubrics with the MR-Bi.

Conventional rubrics	MR-Bi
Cover a single assignment or task during or after a course	Covers a full curriculum across the entire academic career span
Itemise specific assignment elements for grading (layout, figures, discussion, grammar...)	Itemises discipline- and scientific-method-related KSAs for evaluation
Use scores to help grade performance on each element	Uses qualitative descriptions to help assess performance of KSAs
Help learners to identify and understand what they must do to achieve a certain grade	Helps learners to identify their developmental stage and hence to pinpoint their training needs

3.1 Structure of the MR-Bi

So, what does the MR-Bi look like? The axes of the MR-Bi list i) the KSAs that form the bedrock of bioinformatics as a scientific discipline, and that are hence the focus of instruction; and ii) five stages of a developmental trajectory, from less to more expert (*i.e.*, from a student, new to the field, to an experienced, fully independent scientist). For each KSA, the table's cells give brief descriptions (the Performance Level Descriptors (PLDs)) of how a learner might be expected to perform at each stage and thence to change over time.

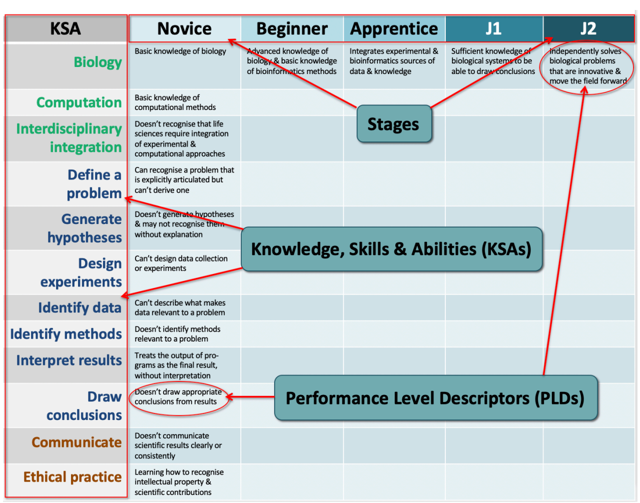


Figure 1 ***Structure of the MR-Bi.*** *The x-axis outlines five stages of a developmental trajectory, from Novice to Journeyman; the y-axis lists the KSAs to be delivered by a course; the cells describe how a learner might typically perform, & change over time, when traversing the trajectory.*

Figure 1 provides a summary of this general structure. Note that, as the figure is an overview, the PLD excerpts shown aren't intended either to be complete or to be fully legible – the entire MR-Bi, including the complete set of PLDs[19], is presented later, in **Table 2**. Let's take a closer look at the MR-Bi's components in turn.

The KSAs

The MR-Bi encapsulates 12 content-/topic-agnostic KSAs. The first two are foundational, discipline-specific KSAs, the rudimentary components of bioinformatics, while the third concerns their integration. Seven other KSAs are based on core elements of the scientific method. The last two reflect the importance of being able to communicate scientific results, and the necessity for sound ethics to underpin all aspects of the scientific enterprise. The full set of KSAs is as follows:

1. *Prerequisite knowledge, biology*
2. *Prerequisite knowledge, computational methods*
3. *Interdisciplinary integration*
4. *Define a problem based on critical review of existing knowledge*
5. *Hypothesis generation*
6. *Experimental design*
7. *Identify data relevant to the problem*
8. *Identify and use appropriate analytical methods*
9. *Interpretation of results/output*
10. *Draw and contextualise conclusions*
11. *Communication*
12. *Ethical Practice*

As the core KSAs relate to generic scientific practice, this structure can be readily customised by changing the discipline-specific KSAs.

A closer look at the MR-Bi's KSAs

The KSAs deliberately aren't restricted to specific facts or content, as tools and topics change quickly. They're broader, higher-level concepts than competencies or KSAs captured in other frameworks, which may drill down, say, to specific scripting skills, programming languages, operating systems, *etc.* The MR-Bi's KSAs provide focal points for building curricula (see table below), making it adaptable and robust.

MR-Bi KSAs	
Prerequisite knowledge (PK), biology and ***PK, computational methods***	Cover the foundational, background knowledge, and basic skills and abilities of biology and computing
Interdisciplinary integration	Concerns the ability to integrate across the bio- and computing domains, and/or other domains
Define a problem based on critical review of existing knowledge	Concerns the application of critical evaluation skills and judgement; aims to promote the ability to identify and solve biological problems
Hypothesis generation and ***Experimental design***	Cover scientific reasoning, statistics, hypothesis testing, methodology and pilot testing
Identify data relevant to the problem and ***Identify & use appropriate analytical methods***	Cover the ability to find and use relevant data and methods, and to understand their strengths and weaknesses
Interpretation of results/output and ***Draw & contextualise conclusions***	Concern the ability to correctly interpret *p*-values and dependencies of multiple methods, and to align results with conclusions and existing knowlege
Communication and ***Ethical practice***	Cover the ability to present scientific work to diverse audiences, and follow ethical practices relating to transparency, rigour and reproducibility

EXERCISES

1 Write down & explain the three principal components of the MR-Bi.
2 Consider these competencies: *General biology*; *Bioinformatics tools & their use*; *Web-based computing skills*; *Command-line skills*; *Professional, ethical, legal & social issues of bioinformatics data*; *Communication of bioinformatics topics to a range of audiences*[10]. Which KSAs (refer to box above) might include each competence?
3 Does any competence span more than one KSA? What challenges might this bring to course designers and/or instructors?

The developmental stages

The Mastery Rubric[14] builds on the European Guild Structure, which outlines a trajectory from Apprentice, through what's known as the Journeyman stage, ultimately to Master Craftsman (or Master Tradesman) status. The MR-Bi differs from this structure by adding Novice and Beginner stages; it further distinguishes itself by eliminating the Master stage and differentiating the Journeyman period into early and late stages (designated J1 and J2), as this is generally the longest phase of training, and qualitative differences are evident between a newly qualified Journeyman and one with, say, 10 or more years of experience. Overall, then, the MR-Bi's trajectory progresses from Novice, Beginner and Apprentice, to the proficient J1 Journeyman, ultimately to the expert, fully independent J2 Journeyman, who is deemed *subject master* and *teacher of the next generation(s)*.

A closer look at the MR-Bi's developmental stages

The stages of the MR-Bi form an evidence-based developmental trajectory of increasing cognitive complexity. Learners' performance, behaviours, habits of mind and required level of supervision at each stage are observably different on the road to independent practice and subject mastery, as outlined in the table below.

Novice	Deals with *facts*: memorises them, generally without questioning; can engage with well-defined problems, with known solutions (*e.g.*, early undergraduate-level thinking)
Beginner	Beginning to understand the uncertainty of scientific 'facts'; uses and applies given tools as instructed (*e.g.*, early Master's-level thinking)
Apprentice	Choooses and applies techniques to given problems; analyses and contextualises results; seeks guidance to improve (*e.g.*, early PhD level)
Journeyman 1	Newly qualified for independent practice; typically still requires some supervision to help evaluate research results (*e.g.*, postdoc level)
Journeyman 2	Fully independent scientific practitioner; expertly analyses, synthesies and evaluates research results (*e.g.*, principal investigator level)

REFLECTIONS

1 Think of a course you teach or that you're currently designing.
2 Considering the table above, can you identify, and explain, the entry- and exit-level developmental stage(s) your course targets?
3 Overall, how many developmental stages does your course span?

The Performance Level Descriptors (PLDs)

For each KSA, at each developmental stage, the MR-Bi provides a set of PLDs, describing performance and mapping progression as learners traverse the trajectory from Novice to Journeyman, gaining greater expertise at each level. The PLDs were devised to give a *broad, high-level guide*, to illustrate the types of learner performance, behaviour or habits of mind that are characteristic at each stage: they generally state what learners *can* do; however, for comparison, they sometimes point to what they can't *yet* do. It's important to note that the PLDs aren't a gold standard of truth, devised by worldwide consensus: they aren't set in stone or intended to be definitive; hence, if deemed more appropriate, instructors could devise other, different (and perhaps more detailsed) PLDs, for example to better reflect the nature of their own courses or programmes. The PLDs here are intended as a starting point: the idea is that they should be familiar as *general* traits, showing instructors how learner performance *typically* changes as their cognitive skills develop over time.

By way of example, a Novice is described, here, as someone who has basic knowledge, who reads and generally understands – but doesn't question – research results, whose thinking is based on uncritical acceptance of given information as factual or true. An Apprentice is beginning to understand the relative strengths of experimental methods, and *does* appreciate the uncertainty in research results, but still requires some guidance. By contrast, the J2 Journeyman is a fully independent expert in design and critical evaluation of experimental methods and their results, can solve innovative biological problems and can generalise to other systems.

A closer look at the MR-Bi's PLDs

PLDs are examples of concrete, observable learner behaviours that, with practice, can be developed over time. They prompt instructors to consider what specific learner behaviours will demonstrate particular KSAs, and what tasks will elicit these behaviours. They thus clarify what instructors need to teach *and* assess at each developmental stage; they also indicate to students what evidence they need to show that they've 'achieved' a given KSA. The table below illustrates how, for KSA *Identify data relevant to the problem*, the PLDs, hence learner behaviours, cognitive skills and independence, change at each stage.

Example PLDs, showing independence evolving across stages	
Novice	Uses data, as directed. Doesn't find relevant data; can't describe what makes data or a given data-resource 'relevant' to a given problem
Apprentice	Can search for data and will ask if unsure about the relevance to a given problem. Learning how to identify (and evaluate strengths/weaknesses of) data-resources, to determine their relevance for a given problem. With guidance, learning how to use these to address given research problems
Journeyman 2	Identifies data that are directly relevant to a problem of own or others' devising. Consistently identifies (and evaluates strengths/weaknesses of) a variety of data-resources that can address a problem or help to formulate it more clearly; recognises if the necessary data don't yet exist

REFLECTIONS

1 Review the PLDs for KSA *Identify data relevant to the problem* in the table above – refer to **Table 2** for more details.
2 In what ways do you think the Novice changes en route to Apprentice level in terms of independent thought and practice?
3 Consider the same question for KSA *PK, computational methods.*

3.2 Bloom's taxonomy in the context of the MR-Bi

Although the concepts of Novice, Beginner and Apprentice may be broadly familiar to many readers, it's likely that the concept of

Journeyman is not. To try to make these developmental stages more concrete in terms of what they mean, or how they relate to cognitive complexity, or academic stage, each can be mapped to a specific level, or levels, of Bloom's taxonomy[21].

Bloom's taxonomy of cognitive complexity

Created in 1956, Bloom's taxonomy is a widely used classification of cognitive skills. It features a six-level hierarchy of increasing complexity, ranging from the basic skill of *remembering* (being able to recall facts and basic concepts) to the advanced skill of *evaluating* (being able to defend opinions or decisions). There has been some debate about the order of the final two levels of the hierarchy (does *synthesising* or does *evaluating* represent the pinnacle of cognitive skills?)[22]; however, leaving the minutiae aside for the sake of simplicity, the original hierarchy is illustrated in **Figure 2**.

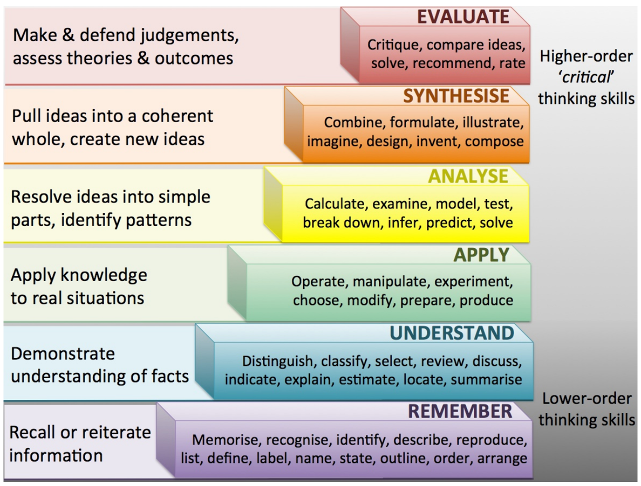


Figure 2 ***The six-level hierarchy of Bloom's taxonomy of cognitive complexity.*** *Remembering information sits at the bottom & evaluation at the top of the hierarchy of thinking skills. Associated with each level is a set of verbs (only a sample is shown here) that express observable & hence measurable learner behaviours characteristic of that level.*

As shown in the Figure, each Bloom's level is accompanied by a set of active verbs that express expected, measurable learner behaviours at that level: *e.g.*, achieving the level *understand* means to be able to classify, select or explain a piece of information: here, *classify*, *select*, *explain* are observable learner behaviours that may be readily assessed by an instructor.

Typical illustrations of the taxonomy, like **Figure 2**, depict successive discrete levels, suggesting a fixed, step-wise *developmental trajectory* from lower- to higher-order critical-thinking skills. However, as mentioned above, an alternative version[22] places *synthesise* (the ability to create new or original work) at the top of the hierarchy; the structure shouldn't therefore be regarded as completely rigid. It's helpful instead to regard the taxonomy as a continuum of cognitive levels (hence the spectral colours used in **Figures 2** and **3**, and in **Table 1**), where each merges into the next, providing a structured tool in which cognitive complexity is made explicit through a set of observable, assessable learner behaviours. Indeed, some qualification frameworks, such as the Dublin Descriptors, elide successive Bloom's levels (*Knowledge and understanding, Applying knowledge and understanding, etc.*) to clarify or simplify the relationships between them[23].

Bloom's, the MR-Bi stages & academic progression

Because the MR-Bi explicitly outlines a developmental trajectory (albeit with five stages rather than six), it's relatively straightforward to relate its stages to Bloom's levels. Moreover, as Bloom's is widely used in the development of education programmes, we can consider how the stages of the MR-Bi, in tandem with Bloom's levels, might relate to traditional stages of academic progression – **Table 1**.

Table 1 ***Relationship between stages of the MR-Bi, Bloom's cognitive levels & academic stages.*** *Some typical characteristics of learners at each stage are described in the right-hand column.*

MR-Bi stage	Bloom's level	Academic stage	Typical learner traits at this level
Late Journeyman (J2)	**late 6:** evaluate	Career **post**doc, PI	**Independent scientist**, expert in design/critical evaluation of experimental paradigms and their results; expertly integrates bioinformatics into research practice; can apply/develop new methods, formulate problems.
Early Journeyman (J1)	**5, early 6:** evaluate, synthesise	Late PhD student, early postdoc	**Proficient scientist**, but still needs mentoring; synthesises knowledge; beginning to critically evaluate experimental paradigms and their results; accepts uncertainty; contributes to problem formulation.
Apprentice	**3-4, early 5:** synthesise, analyse, apply	Master's early PhD student	**Fluent scientist,** who can choose and apply methods to given problems, analyse and interpret data, identify basic limitations, and contextualise results; doesn't generate new problems; seeks guidance to improve performance.
Beginner	**2-3:** apply, understand	Late undergraduate, early Master's	**Learning how** to analyse given problems; beginning to understand uncertainty; can use tools and apply them as instructed.
Novice	**1-2:** understand, remember	Early undergraduate	**Engages with** given problems, with known solutions, but doesn't question research results; limited understanding of uncertainty.

REFLECTIONS

1 Consider **Table 1**. In what ways do you think the Novice changes en route to Apprentice level in terms of critical thinking?

2 Are these changes evident within the PLDs for KSAs *Identify data relevant to the problem* and *PK, computational methods?*

We can now frame these basic elements within the complete MR-Bi, which is presented in **Table 2**. The structure is the same as that in **Figure 1**, but includes a general description (broadly tracking the *typical* development of a student progressing through university) in the first row, and a description of the requisite Bloom's level in the second; these are followed by the KSAs and PLDs at each stage.

Table 2 ***The complete Mastery Rubric for Bioinformatics (MR-Bi).*** *PLDs are shown at each stage for each KSA; they are preceded by general descriptions of a bioinformatics practitioner at each stage & a set of considerations for evidence of performance at the requisite Bloom's level.*

	Novice	Beginner	Apprentice	J1 Journeyman	J2 Journeyman
General description of a bioinformatics practitioner	Reads, generally understands, but doesn't question, life science research results. Beginning to recognise that 'facts' are actually just the best-currently-supported theory. Limited engagement with uncertainty associated with 'facts'; developing understanding of experimental design paradigms in biology, & own specific area of study.	Consolidates reading & understanding, beginning to learn how to analyse given biology problems (with software). Growing recognition that 'facts' are typically the best-currently-supported theory. Engaging consistently with uncertainty associated with 'facts'; deepening understanding of experimental design paradigms in biology, & own specific area of study.	Reads & understands; reliably identifies methods (software & programming) for given problems. Chooses & executes correct analysis, but not necessarily able to identify several methods that could be equally viable, depending on given research objectives. Qualified as a fluent, but not as an independent, scientist who uses bioinformatics as a tool, but doesn't yet synthesise technology with biology to generate new research problems.	Qualified as a proficient, independent scientist who uses bioinformatics methodologies as part of routine practice. Can pose novel scientific questions; identifies data & technology to align appropriate statistical/analytical methods to desired scientific objectives. Experienced reviewer of technical features of bioinformatics methods. Newly independent, able to integrate bioinformatics techniques into novel research problems in area of expertise.	Independent scientist who expertly integrates bioinformatics & more traditional methodologies, as needed, to achieve desired objectives & contribute to the body of knowledge. Expert reviewer of relevant technical features of available bioinformatics options.
Considerations for evidence of performance at this level	**Bloom's 1, early 2: *remember, understand*.** Can engage with well-defined problems, with known solutions. Work doesn't generally reflect self-assessment.	**Bloom's 2-3: *understand, apply*.** Can engage with well-defined problems. Applies *only what he/she is told to apply*. Work reflects some self-assessment, when directed to do so.	**Bloom's 3-4, early 5: *apply, analyse, synthesise*.** Can choose & apply techniques to problems that have been defined (by or with others). Can analyse & interpret appropriate data, identify basic limitations & conceptualise a need for next steps, & for contextualisation of results with the literature. Seeks guidance to improve self-assessment.	**Bloom's 5, early 6: *evaluate, synthesise*.** Evaluates life-science knowledge, while developing abilities to integrate bioinformatics into research practice. Shows independent expertise in a specific life-science area, & confidently integrates bioinformatics technology into that area. Beginning to critically evaluate experimental paradigms & their results, without requiring there to be 'one right answer'. Consistently self-assesses.	**Bloom's 6: *evaluate*.** Prepared for independent scientific work. Expert in design & critical evaluation of experimental paradigms & their results. Consistently self-assesses, & encourages others to develop this skill.
Ethical practice	Exhibits respect for community standards/ rules for public behaviour & personal interaction. Learning to recognise, & show respect for, intellectual property, professional accountability & scientific contributions.	Learning to recognise scientific 'misconduct' . Learning to avoid, & respond to, misconduct, & the importance of neither condoning nor promoting it.	Learning the principles of ethical professional & scientific conduct. Seeks guidance to strengthen applications of these principles in own practice. Learning how to respond to unethical practice.	Practices bioinformatics in an ethical way, & doesn't promote or tolerate professional or scientific misconduct. Seeks guidance in how/when to take appropriate action when aware of unethical practices by others.	Practices, & encourages all others to practice, bioinformatics in an ethical way. Doesn't promote or tolerate professional or scientific misconduct. Takes appropriate action when aware of unethical practices by others.
Prerequisite knowledge – biology (includes statistical inference & experimental design considerations)	Basic knowledge of biology; little-to-no awareness of the uncertainty inherent in experimental designs common in the life sciences. Thinking about the life sciences is based on uncritical acceptance of information as 'factual' or 'true'.	Advanced knowledge of biology, & basic knowledge of key bioinformatics methods. Can run very simple statistics/programs to answer pre-defined scientific questions. Learning to understand the uncertainty inherent in the scientific method; questions assumptions in the data & their relevance for given scientific problems (which are defined by others).	Integrates experimental & bioinformatics/technological sources of data & knowledge. Understands the uncertainty inherent in the scientific method; questions assumptions in the data & their relevance for given scientific problems (which are typically defined by or with others). Exploits experimental design & statistical inference, with guidance, to answer given scientific problems. Recognises inconsistencies in biological data/experiments identified by others, but can't troubleshoot experimental methods independently.	Recognises the importance of, & is able to critically evaluate, the relevant literature, & understands historical background of the relevant biological system(s). Sufficient knowledge of a biological system(s) to be able to draw functional conclusions from analytical results. Collaborates with experts to inform the next stages in the experimental design process (validating results, follow-up analyses, *etc.*).	Makes predictions to inform next stages of the experimental design process. Evaluates relevant experimental methods that can be applied in any problem. Can generalise to other biological systems; independently solves biological problems that are innovative & move the field forward.

Prerequisite knowledge – computational methods (includes statistical inference & experimental design considerations)	Basic knowledge of computational methods; little-to-no awareness of the relevance of computational methods for life sciences. No awareness of experimental designs or how these can be used or implemented in computational applications. Thinking about tools, computers, software & programming is strictly uni-dimensional: *i.e.*, extrapolation &/or abstraction of knowledge about computational methods to other systems, programs or problems aren't possible. Can run given software or execute given code with precise instructions; can't write a script or debug/troubleshoot.	Computers, software, tools & programming are understood to be options for scientific work. Learning how to write & test code, run software, or use tools, as appropriate. Developing awareness of the variety of bioinformatics tools, designs & resources, but isn't able to choose or apply the most appropriate of these for any given question; when choices are made, tools are used uncritically. Developing awareness that computational tools require input parameters, but uses default settings. Learning to read, understand, troubleshoot & make minor modifications to existing code/ scripts. Doesn't synthesise results or outputs.	Learning to test software & programming approaches to different types of problem. Experimental design & statistical inference using computing & algorithms are recognised & applied, with guidance, to answer given scientific problems. Learning best practices for programming, if programming is part of the task. Can write basic code in a given language or run appropriate software, using judgement, but not inventing or innovating. Can't troubleshoot complex computational methods – will ask for guidance. Exploring alternatives to default input parameters across computational tools. Can apply knowledge of tools to interpret results & output. Seeks guidance in synthesis of results or outputs.	Recognises the importance of, & critically evaluates & understands, historical background of the relevant data, databases, algorithms, tools, analysis/statistical methods & computational resources. Can use these & justify trade-offs across methodologies (*e.g.*, which statistical test to apply & what computational methods to use). Collaboratively synthesises & critically questions analysis results & output from tools. Recognises the iterative nature of experiments (*e.g.*, bench, data analysis, back to bench). Can write code/use tools to accomplish these, but collaborates with experts for identifying & articulating biological problems that are innovative & move the field forward.	Develops robust, well-documented, optimised, reproducible code &/or uses tools to address biological problems; moves away from standard procedures & innovates to accommodate new data types, tools & techniques, as needed. Can generalise to new coding languages or software/tools/resources.
Integrate interdisciplinarity	Doesn't recognise life sciences as requiring integration of both experimental & computational/modelling approaches. Perceives disciplines as separate; integration only occurs when/as directed. Information, ideas & tools that are interdisciplinary are used without question.	Beginning to think about life sciences as requiring integration of experimental & computational/modelling approaches. Recognises that interdisciplinarity is needed, but doesn't know how (or when) to do it, & requires direction. Learning the integrating process; learning strengths & weaknesses of biological & computational methods, but not sufficient to question assumptions from these & other disciplines.	Understands that life sciences integrate both experimental & computational/modelling approaches; seeks guidance about how & when to integrate. Developing an understanding of the strengths & weaknesses of biological & computational methods, beginning to question fundamental assumptions from these & other disciplines for any given scientific problem (which is typically defined by, or in conjunction with, others).	Collaboratively integrates across relevant disciplines to address, & solve, innovative biological problems. Tests multiple avenues to triangulate solutions, with minimal guidance. Recognises the roles of interdisciplinary teams in the research process, & the importance of integrating interdisciplinarity early on. Works effectively on interdisciplinary teams with minimal guidance.	Formulates innovative biological problems that require interdisciplinary solutions. Integrates methods & results to derive & contextualise solutions to biological problems. Consistently tests multiple avenues to triangulate solutions, while exploiting relevant findings from other disciplines. Actively builds interdisciplinary teams, as needed.
Define a problem based on a critical review of existing knowledge	Can recognise a problem that's explicitly articulated or concretely given, but can't derive one. Unaware of the depth & breadth of the knowledge base that is, or could be, relevant for the formulation of a problem. Doesn't recognise design features or other evidence as the basis of/support for problem articulation. Doesn't recognise uncertainty or how this affects the formulation of solveable problems.	Developing awareness of the depth & breadth of the knowledge base that is, or could be, relevant for the formulation of a problem. Can't differentiate gaps in own knowledge from gaps in 'the knowledge base'. Developing the ability to recognise that uncertainty may have arisen in the formulation of solutions to problems.	Beginning to use, with guidance, the appropriate knowledge base to address a given problem. Recognises the need to consider a wider scope of knowledge for alternative solutions to a problem common across contexts or domains. In guided critical reviews, learning to recognise that design features & evidence base are important to drawing conclusions. Recognises the role of uncertainty in research, & that reproducibility & potential bias should be considered for every result.	Can explore & critically review the relevant knowledge base, & collaboratively articulate a problem based on that review. Reviews include assessment of relevance from (potentially) ancillary domains, bias, reproducibility & rigour; recognises when appropriate & inappropriate methodology is used. Recognises when incomplete review is provided (by themselves or others). Can discern reproducible from non-reproducible results; can identify major sources of bias in the knowledge base.	Independently defines & articulates theoretical or methodological problems based on a critical review of the relevant knowledge base(s). Knows the hallmarks of questionable research hypotheses & misalignment of testing/statistics with poorly articulated research problems; consistently finds & identifies sources of bias. Articulates when appropriate & inappropriate methodology is used/reported. Critical review & problem articulation integrate diverse disciplinary perspectives when appropriate.
Hypothesis generation	When directed, follows instructions to *test* hypotheses; doesn't generate them & may not recognise them without explanation. Uses the default settings of software & other tools, rather than a hypothesis, to guide any analysis. Doesn't question methods to be used, or	When directed, uses the default settings of software, tools or GUI to test hypotheses in pre-planned analyses; doesn't generate testable hypotheses. Doesn't recognise that hypotheses may be generated & tested within the intermediate steps of an analysis. Developing the	With guidance, can leverage tools, software, data & other technologies (GUI/programming) to test hypotheses; can generate hypotheses based on the data or the technology, but not on their combination. Hypothesis generation possible in concrete, fully parameterised problems; developing the ability to	Collaboratively integrates hypothesis generation into the consideration of literature, data & analysis options. Seeks appropriate guidance in the synthesis of data & technology to generate novel, testable hypotheses. Considers the process of hypothesis generation & testing to be iterative, when this is	Independently generates testable hypotheses that are scientifically innovative as well as feasible (possible for economic reasons, time, impact, *etc.*). In own & others' work, recognises that, & articulates how, hypothesis generation from planned & unplanned analyses differ in their evidentiary weight & their need for independent

	assumptions of methods that are used.	understanding that all methods involve assumptions.	identify whether a hypothesis is testable. Learning to recognise that experimental design & design of software/programming solutions include hypothesis generation to some extent. Developing the abilities to identify, & plan to address, assumptions that different hypotheses necessitate.	appropriate. Hypothesis generation is done with consideration of reproducibility & potential for bias, & takes into account the most clearly relevant literature; recognises that less-obviously relevant literature may also be informative for hypothesis generation.	replication. Fully explores all relevant knowledge base(s) to support rigour & reproducibility, & to avoid bias, in the generation of hypotheses.
Experimental design	Can recognise concrete features of experiments only if they're described/given and they match basic design elements (*e.g.*, dependent, independent variables). Can't design data collection or experiments. Unaware of covariates or their importance in analysis or interpretation. Doesn't recognise the importance of design, data collection, data quality, storage/access, analysis & interpretation to promote rigour & reproducibility in experimental design.	Can identify salient features of experiments that are dscribed/given, if they match previously encountered design elements, but can't derive them if they're not present. Recognises covariates if mentioned, but doesn't require formal consideration (or justification) or evaluation of covariates. Doesn't recognise that one experiment alone can't adequately address meaningful biological research problems. Understands that experimental design involves identifying, gathering, storing, analysing, interpreting & integrating data & results.	Can match the correct data-collection design to the instruments & outcomes of interest. May include/exclude covariates, or other design features, 'because that is what's done', without being able to justify their roles in the hypotheses to be tested. Developing an understanding that weak experimental design yields weak data & results. Needs help in conceptualising covariates & their potential roles in planned analyses. Beginning to recognise that, & can explain why, one study is usually insufficient to answer given research problems/solve biological problems adequately. Follows templates for identifying, gathering, storing, analysing, interpreting & integrating data. Learning to consider reproducibility & rigour in experimental design, & to question templates that do/don't include these concepts.	Recognising that explicit attention to experimental design will result in more informative data; designs experiments in consultation with experts in content & statistics: these experiments may include power calculation considerations, if relevant; modelling requirements; measurement/sampling error & missing data. Collaboratively designs experiments that address the need for reproducibility & sensitivity analysis. Learning to conceptualise pilot studies & sensitivity analyses. Learning to adapt problems so that hypotheses can be generated & made testable via experiments.	Independently designs appropriate & reproducible experiments & other data-collection projects, using methodologies that are aligned with the testing of specific hypotheses. Consistently identifies & justifies choices of instruments & outcomes (& covariates if relevant). Collaborates with experts as needed on appropriate use of advanced methods, including accommodating measurement & sampling error, attrition (if needed) & modelling requirements; can adapt complex problems so that hypotheses can be generated & made testable via experiments. Understands, & can exploit the strengths & weaknesses of, experimental design, data & modelling approaches with respect to the biological problem under consideration. Uses pilot studies & sensitivity analyses appropriately.
Identify data relevant to the problem	Uses data, as directed. Doesn't find relevant data; can't describe what makes data or a given data-resource relevant to a given problem.	Correctly uses data that are provided, or can follow a script/'recipe' to obtain (access, manage) relevant data to which they're guided. Can't determine whether a given data-set or -type is relevant for a given problem, but is developing an awareness that not all data are equally relevant, or equally well suited, to all research problems. Developing awareness of the features of data/data-resources that constitute 'relevance', & that these features must be assessed before choosing data to use.	Can initiate a search for data & will ask if uncertain about the relevance for any given problem. Learning how to identify, & evaluate strengths & weaknesses of, data-resources, to determine whether a given data-set or data-type is relevant for a given problem; &, with guidance, learning how to leverage these to address given research problems. Learning how reproducibility can be affected by the choice (& features) of data.	Collaboratively identifies relevant data-resources. Understands the relative strengths & weaknesses of data-sets & -types for addressing a specific problem. Learning to address & formulate scientific problems (based on recognised gaps in the knowledge base) using relevant data-resources. In own & others' work, recognises that, & articulates how, choices for data (collection or use) require assumptions & justification, & must yield reproducible results.	Identifies data that are relevant to a problem. Consistently identifies, & evaluates strengths & weaknesses of, data-resources that can address a problem or help to formulate it more clearly; recognises if the necessary data don't yet exist. Justifies the relevance of data-sets to a problem in terms of their individual strengths & weaknesses. Articulates hypotheses, & designs experiments, that leverage strengths in the data; includes triangulating data or results to address weaknesses in the data. Identifies whether data appropriate to the specific scientific question were used when reviewing proposals, protocols, manuscripts &/or other documentation describing data & research results.

Identify & use appropriate analytical methods	Uses methods that are provided & in a given order (*i.e.*, a pipeline; & treats workflows* as if they're pipelines). Doesn't identify relevant methods; can't describe what makes a method relevant to a given problem. Unaware that methods & software have default settings. Doesn't question propriety, assumptions or the order of methods employed; focus is on the superficial attributes of given methods & protocols.	Uses methods as directed, & learning about the concepts of pipelines & workflows; still uses workflows as if they're pipelines, but beginning to attend to decision points. Learning to recognise pros & cons of methods/software, but can't yet effectively compare, evaluate or rank them. Becoming aware of default settings of software or methods & their effects on results; & beginning to explore & enquire about tailored settings. Understands that more than one method/tool may be available to deal with a problem or data-type, but can't choose effectively. Learning about similarities & differences across methods, & that choices (particularly of multiple methodologies for one question) should leverage independence of methods to support reproducible results.	Can identify methods, software & pipelines that are relevant for a given problem; seeks guidance about the best approach. Learning to evaluate/rank & justify alternative methods in terms of general features of their efficiency & relevance for the given research problem. Beginning to recognise that a 'pipeline' involves only the choice of which method(s) to use; while a 'workflow' requires many choices & decisions. With guidance, seeks to identify & implement appropriate workflows to address given research problems. Learning how reproducibility can be affected by the choice & implementation of methods, including independent replication of essentially the same method vs. independent replication using diverse methods.	Collaboratively considers the knowledge base, & features of the relevant data & analysis options, in identifying the most appropriate approach(es) to tackle a scientific question. Uses appropriate analytic methods, pipelines & workflows, recognising, & taking advantage of the fact, that these may represent distinct approaches to the same problem. Knows when & how to control False Discovery Rates (FDR) to promote reproducible results across methods. In own & others' work, recognises that, & articulates how, choices for methods, pipelines & workflows require assumptions & justification, & must yield reproducible results.	Recognises if/when the necessary methods, pipelines & workflows to tackle a scientific question *don't yet* exist. Consistently controls FDR to promote reproducible results. Identifies whether appropriate analytical methods were used when reviewing proposals, protocols, manuscripts &/or other documentation describing methods, pipelines, workflows & research results.
Interpretation of results/output	Treats the output of a program as the final/complete result – with no interpretation required – & is unaware of the concepts of validation & cross-validation or their importance for interpretation of results/output. Uses the *p*-value to indicate 'truth' in statistical analysis. Over-interpretation is typical. Unaware of the importance of FDR controls. Doesn't seek coherence in/recognise incoherence of results with the analysis plan or pipeline; can't align methods, results & interpretation.	Interpretation of results depends on *p*-values, but understanding of *p*-values is incomplete. Learning to recognise that interpretation of output critically depends on methods used & the pipeline in which the results are obtained. Developing awareness of FDR controls. Learning that the interpretation of immediate results could be an interim step in an overall problem-solving context.	Seeks guidance to interpret results/output, including considerations of alignment of methods & results. Understands that the *p*-value represents evidence about the null hypothesis, not the study hypothesis, but doesn't consistently avoid reification. Recognises that, but doesn't always act as if, very small *p*-values are *not* 'highly significant results'. Can apply FDR controls, but does so only when reminded/required. Recognises when the interpretation of immediate results is an interim step in an overall problem-solving context.	Can discern, based on results, methods & hypotheses, whether more experiments &/or data-processing are required for robust result interpretation; collaboratively uses the appropriate knowledge base & data-resources to interpret results; resists reification & is committed to good-faith efforts to falsify hypotheses. Consistently & appropriately uses FDR controls.	Interprets results critically & with respect to the analysis plan; seeks/promotes alignment of methods, results & interpretation. Prioritises interpretable & reproducible results above any other outcome (*e.g.*, publication or completion of tasks/project), & insists on FDR controls & other sensitivity analyses in all work. Avoids problems that can arise in interpreting results, including bias, reification & other failures of positivism. Evaluates the quality & appropriateness of procedures, statistical analyses & models when reviewing papers & projects/proposals, based on the writers' – & own – interpretation of results.
Draw & contextualise conclusions	Doesn't draw appropriate conclusions from given results; without direction, will not even contextualise conclusions with the protocol that was followed. Not aware of the difference between conclusions about the null hypothesis & those about the research hypothesis. Conclusions may over- or understate results & be driven by *p*-values or other superficial cues. Doesn't recognise the importance of identifying & acknowledging methodological limitations, or their implications, for conclusions.	Learning fundamentals of how appropriate conclusions are drawn from results, but may not be able to draw those conclusions from given results themselves. Learning to differentiate between conclusions about the null hypothesis & those about the research hypothesis. Learning why *p*-value-driven conclusions, & the lack of FDR controls, are not conducive to reproducible work. Conclusions are generally aligned with given results, but when multiple methods are used, doesn't recognise the	With guidance, can draw conclusions in own work that are coherent with the research hypothesis/hypotheses & across the entire manuscript/write-up (as appropriate). Learning to critically contextualise results; draws the most obvious conclusions, but struggles to see patterns, or draw more subtle conclusions. Learning that 'full' contextualisation of conclusions requires consideration of limitations deriving from methods & their applications, & their effects on results & conclusions. Learning to recognise how independence of multiple methods applied to similar	Can extract scientific meaning from data analysis & knows the difference between statistical & biological significance. In own & others' work, seeks competing, plausible alternative conclusions. Can judge the scientific importance of results, & draw conclusions accordingly. Can draw conclusions & contextualise results with respect to an entire manuscript/write-up in a given project or study, or to the literature (as appropriate). Can detect when conclusions aren't aligned with other aspects of the work (*e.g.*, introduction/background,	Expertly contextualises results & conclusions with prior literature, across experiments or studies, & within any given document (*e.g.*, manuscript, write-up). Strives to fully contextualise conclusions in own work, & also requires this in others' work. Draws & contextualises more subtle conclusions than at earlier stages. Can conceptualise new experiments based on the lack of robust &/or defensible conclusions in others' work. Carefully considers consistency of conclusions with the other parts of own or others' work.

	Doesn't, or can't, apply rules of logic to scientific arguments, & commits logical fallacies when drawing conclusions.	dependencies among methods that appear to reinforce, but actually replicate, results. Conclusions are neither fully contextualised with the rest of a document (write-up, paper, *etc.*) or study/experiments/paradigm (contextualisation for *coherence*), nor with the literature (*critical contextualisation*).	data/problems supports reproducible conclusions.	methods, results). Gives careful consideration to limitations deriving from the method & its application in a specific study. Sees patterns, & perceives more subtle conclusions than earlier-stage scientists, & collaborates to fully articulate & motivate them. Writes the Discussion & Conclusions sections, including limitations, of own articles, with collaboration.	
Communication	Doesn't communicate scientific information clearly or consistently; is unaware of community standards for scientific communication. Generally relies on lay summaries to support own communication; doesn't recognise that using original literature strengthens scientific communication. Doesn't differentiate appropriate & inappropriate scientific communication, nor understand the ethical implications of each.	Learning both to recognise the value of clear communication, & about the role of communication in sharing & publishing research, data, code, data-management protocols, tools & resources. Developing an awareness of community standards for scientific communication, & that these include documenting code, annotating data & adding appropriate metadata. Doesn't adapt communication to fit the receiver. Learning to differentiate appropriate & inappropriate scientific communication, but doesn't yet understand that transparency in all communication represents ethical practice, even when the desired results have not been achieved.	Understands the roles of sharing & publishing research, data, code, data-management protocols, tools & resources in scientific communication. Seeks guidance so that own communication is coherent, accurate & consistent with community standards (*e.g.*, following FAIR‡ principles; ensuring socially responsible science). Learning to document code, annotate data & add appropriate metadata – & the importance of these (as appropriate given their research/context) for sharing & integration. Learning the importance of adapting communication to fit the receiver, seeking opportunities to practice this. Learning that transparency in all communication represents ethical practice, even when the desired results have not been achieved.	Consistently & proficiently uses technical language to correctly describe what was done, why & how. Sufficient consideration given to limitations, with explicit contextualisation of results consistently included in the communication of results & their interpretation. Adapts communication to fit the receiver; recognises that sometimes communication must be consistent with community standards beyond own discipline. Appropriately documents/ annotates all data, code, tools & resources for sharing, integration & re-use. Understands that transparency in all communication represents ethical practice.	Expert communicator & reviewer of scientific communication; adheres to, & promotes, disciplinary standards for communication. Communicates in a manner that is consistent with standards across communities beyond own discipline, as appropriate. Ensures communication is appropriate for a target audience, expertly adapting to fit the receiver(s). Communication is transparent, & appropriate to support reproducibility – &, thereby, ethical practice – in every context.

*Workflows support decisions: they aren't necessarily linear, but can be multi-directional and iterative; any point can be iterated, or new starts from within the workflow can be made. Pipelines are uni-directional, not iterative, and don't have decision points. Pipelines can exist within workflows, but workflows don't exist in pipelines.
‡ FAIR: Findable, Accessible, Interoperable, Reusable.

Clearly, there's a lot to digest in **Table 2**, and readers are not expected to assimilate all the details; rather, the MR-Bi should be seen as a multi-functional tool, from which users may select, and focus on, only those parts that are required to achieve a given purpose (more guidance on using the MR-Bi is available in the companion Guide, *Using the Mastery Rubric for Bioinformatics – a Professional Guide*[20]).

For now, without getting too distracted by the details, we can make a few key observations. Note, for example, how the PLDs evolve between stages, how the need for guidance diminishes along the route to independent thought and practice, and how the sense of self-awareness changes: *e.g.*, Novices have gaps in their knowledge, but generally lack awareness of them, while Beginners are starting to recognise that gaps exist – both need guidance; Apprentices do recognise limits to their knowledge, and will actively seek help to try to address them, while Journeymen strategically seek to collaborate with those whose expertise complements their own. These general categorisations, or stages, broadly map to recognisable steps along the academic trajectory, from undergraduate to principal investigator. Note also how the KSAs are heavily influenced by core aspects of the scientific method and scientific reasoning; hence, the MR-Bi doesn't focus on subject-specific content (R or Ruby programming, using BLAST, ClustalO, Galaxy, *etc.*), but does seek to move learners towards independent scientific practice.

The overall structure of the MR-Bi is summarised in **Figure 3**. The figure illustrates how each developmental stage builds**,** layer upon layer, onto the next in terms of cognitive complexity (advancing from Bloom's level 1 to level 6) as a learner progresses from less to more expert, from Novice (outermost layer) to independent scientist (innermost layer). Beneath each layer (not shown) are the PLDs that describe learner performance for each KSA at each level, as detailed in **Table 2**. Together, the KSAs and their PLDs promote scientific problem formulation and problem solving, lending the MR-Bi durability and flexibility.

EXERCISES

1 Consider your own level of bioinformatics training. For each KSA, write down the stage that most appropriately reflects your level of expertise (reviewing the PLDs in **Table 2** might help you do this)**.**

2 Alternatively, examine the high-level summary shown in **Figure 3**. On the figure, for each KSA, tick the stage with which you identify.

3 Are there KSAs in which you are less proficient? If so, can you pinpoint the type of training or practice that might help you progress to a higher level of accomplishment for that, or those, KSA(s)?

4 If you supervise Master's or PhD students, how might this approach be used to identify their training needs?

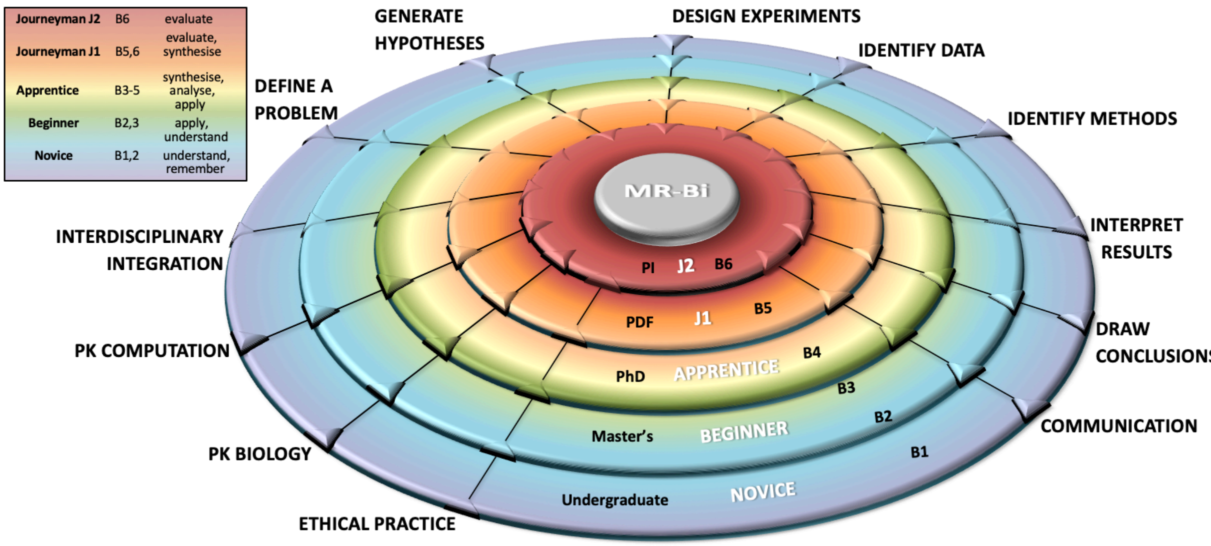


Figure 3 ***Overview of the structure of the MR-Bi.*** *12 KSAs (outermost labels) are encapsulated, together with a developmental trajectory, from Novice (outer circle) to late Journeyman (inner circle). The stages broadly map to the continuum of Bloom's levels, B1-B6 (denoted by the blending of each spectral colour into the next) & to familiar stages of academic progression, from undergraduate, Master's & PhD (which require considerable supervision) to the increasingly independent Post Doctoral Fellow (PDF), & thereafter to Principal Investigator (PI) & subject mastery. The structure is highly adaptable to other disciplines because only two of its KSAs are discipline-specific, so these alone would need altering to focus the Rubric on closely related subjects.*

4 Discussion

A Mastery Rubric is an organising framework that both articulates KSAs, and describes and *stages* their performance levels such that they can be achieved progressively. As we've seen, the framework has three components: i) a set of domain-relevant, transferable KSAs; ii) a set of developmental stages denoting progression along a path of increasing cognitive complexity, towards independence; and iii) descriptions of the range of expected KSA performance levels.

The interplay between these elements affords the Mastery Rubric significant flexibility: it facilitates consistent evaluation of performance of any given KSA, and recognises that individuals may be at different levels in different KSAs, and may progress through them at different speeds. This allows individuals wishing to acquire bioinformatics skills to locate themselves within the table, regardless of their current skill level or disciplinary background: *e.g.*, a person may consider him/herself a *J2 Journeyman in the life sciences*, yet a *Novice in computational methods*. Importantly, then, the MR-Bi can pinpoint a learner's stage (hence current level of performance of any KSA), and also explicitly highlight a route(s) for self-directed learning, from lower-level skills to higher levels of achievement. This feature can also be exploited by instructors, who may have a mix of students within their class with different aptitudes, some at Novice and some at Beginner level of cognitive complexity, with others perhaps even at Apprentice level. Such understanding can help instructors to pitch, and if necessary to adjust, their teaching accordingly. Similarly, it can help mentors to identify, and thence to plug, skills gaps in the doctoral students or post-doctoral researchers under their supervision.

Being built on core scientific-method-related KSAs, the MR-Bi is essentially a standard framework that's readily adaptable to related disciplines simply by changing the discipline-specific KSAs. So, for example, if an instructor wished to develop a Mastery Rubric for a closely related discipline – say, health informatics – the resulting Rubric would have virtually the same KSAs, but with *Prerequisite knowledge of biology* replaced by *Prerequisite knowledge of health sciences*. The PLDs in that Mastery Rubric would then be tailored to describe development of the health-informatics practitioner.

Elaborating any kind of framework to support the development of skills or competencies is challenging: the task involves multiple stakeholders, from different backgrounds, with diverse perspectives and disparate educational goals. One of the challenges is the lack of a standard vocabulary: some frameworks refer to knowledge, skills and *attitudes*[24], others refer to knowledge, skills and *behaviours*[25], while the MR-Bi describes knowledge, skills and *abilities*[19]. In the latter, *abilities* is the preferred term because these are considered more tangible and observable than, say, attitudes. The box below clarifies the distinction between knowledge, skills and abilities.

A closer look at knowledge, skills & abilities

The terms *knowledge*, *skills* and *abilities* can be confusing: indeed, the distinction between skills and abilities can be especially troubling, as these are often used interchangeably. Accepting that many different definitions exist, the simple working guide below can help to understand their meaning in the context of the MR-Bi, where emphasis is placed on what can be reliably observed and hence measured.

Knowledge	The *conceptual* or *theoretical* understanding of facts or information
Skills	The *practical* execution of particular tasks or actions
Abilities	The efficacy with which (*i.e.,* how well) knowledge is put into action or skills are performed, given time, energy, motivation and practice

In addition, as we saw earlier, the MR-Bi's KSAs are much broader, higher-level concepts than the knowledge, skills and abilities/attitudes/behaviours and competencies encapsulated in other frameworks: *e.g.*, Mulder *et al.*[10] define 16 core bioinformatics competencies, including command-line and scripting skills, Web-based computing skills, creating software systems, defining computing requirements, *etc.*; while Matser *et al.*[25] describe ~30 fine-grained competencies, including knowledge of operating systems, writing/adapting computer programs and scripts, parallel programming, installing simulation software, *etc*. Many of these competencies are encapsulated implicitly in the MR-Bi's broader KSAs (here, for example, the 'parent' KSA would be *PK, computational methods*).

Together, the PLDs and KSAs focus on fostering independent scientific practice and developing critical-thinking skills. This emphasis obviates the need either to enumerate all possible subject-specific competencies (the details of which are likely to change over time) or to articulate individual profiles (*personae*) for particular types of practitioner in different settings – in the workplace, individuals will practice bioinformatics in very different ways according to a vast array of possible roles; trying to describe unique *personae* for all such roles is therefore likely to be a relentlessly challenging task (the issues surrounding the robustness, authenticity and scalability of personae in the field of informatics are well-documented[26]).

Nevertheless, understanding the MR-Bi and its applications still requires thought and time. Because it was developed as a tool for *curriculum* development, it may be difficult to see how it can be used to design short training courses. Here, however, it can help instructors to focus on prerequisite knowledge and teaching goals (and requisite learning outcomes) that are time-limited; with its explicit developmental trajectory, it can also be used to direct individuals' acquisition of new, or to deepen existing, skills: *i.e.*, it can help learners recognise their own training needs, identify targeted training opportunities, and thus track their professional development from their current to a higher level of performance.

It's worth making one last point about the MR-Bi's developmental trajectory. At the top of the learning tree is the J2 Journeyman, an independent scientist who, via years of training, has become a discipline expert or subject 'master'. These professionals (like the Guilds' Master Craftsmen) are generally charged with teaching students (the Novices, Beginners and Apprentices) and likely also with mentoring doctoral students and postdocs (the J1 Journeymen). However, many of these individuals will never have been taught *how* to teach: as qualified experts, it's traditionally been assumed (and thus expected) that they're intuitively equipped to convey their mastery to classes full of eager students, or to labs full of enthusiastic researchers.

Recently, the tide has been turning against this assumption, with growing recognition that teaching is itself a skill needing to be taught and nurtured. Reflecting this notion, Mastery Rubrics treat 'subject mastery' separately from the 'Master Level', for which a unique Mastery Rubric for the Master Level (MR-ML) has been created, focusing on teaching and learning about *teaching and learning*[27]. It isn't in scope to discuss the MR-ML here; suffice it to say that it articulates five KSAs (including setting teaching and learning goals, designing learning experiences, and evaluating teaching) at developmental levels Apprentice Master, Journeyman Master and Master. Those wishing to know more are encouraged to read Tractenberg 2021[27].

A closer look at mastery

The concept of 'mastery' has different connotations, according to the context and era in which it's used. The status of master grew from medieval trades and crafts, and was ultimately enshrined in the European Guild structure. Here, an apprentice would learn a trade from a field expert – a master craftsman; having trained with that master for several years, and produced a qualifying piece of work, the apprentice could be recognised as a journeyman. This afforded the individual opportunities to travel across Europe to learn new skills from different masters[28]. After several more years of experience, and often the submission of a 'masterpiece', approved by the Guild masters, journeymen could then be received as master craftsmen[29], thence able to take on – and teach – their own apprentices.

A key outcome of the Guild framework was the creation of universities at Bologna, Oxford and Paris, which began as guilds of students or masters[30]. The Master's degree dates back to those European universities. Then, an individual who'd earned 'mastership' – a master – was allowed to teach in any other university. Since then, the Master's degree (and with it, the notion of master and/or mastery) has changed significantly. Today, the Master's often sits as a kind of stepping-stone between Batchelor's and doctoral degrees; and, despite its name, few would grant their Master's students a licence to teach! Nevertheless, it's long been expected that those progressing beyond PhD level *will* teach, because they've reached the tops of their fields. Today, subject mastery and the ability to teach are recognised as very different skills, the latter itself requiring teaching and nurturing in its own right.

In this Guide, we've seen that the MR-Bi provides a framework for decision-making and learner progression. The ways in which the tool may be used to inform structured approaches to course design is the subject of its companion Guide, *Using the Mastery Rubric for Bioinformatics – a Professional Guide*[20].

TAKE HOMES

1. The MR-Bi maps performance as learners traverse a developmental trajectory from lower- to higher-order critical-thinking skills;
2. The tool can be used to assess learner development and to inform course design; it can also facilitate self-assessment and hence help individuals to recognise their own training needs;
3. The MR-Bi defines five developmental stages (Novice, Beginner, Apprentice, early- and late-Journeyman) on a trajectory to independent practice and subject mastery;
4. The MR-Bi also defines 12 Knowledge, Skills and Abilities (KSAs): two of these are discipline-specific; others are based on core elements of the scientific method and scientific reasoning;
5. For each KSA, associated Performance Level Descriptors (PLDs) describe how performance and critical thinking change as learners progress through the developmental trajectory;
6. The developmental stages can be mapped both to specific Bloom's-level(s) of cognitive complexity and to broad academic levels; and
7. The MR-Bi can be adapted to apply to closely related scientific disciplines simply by changing the discipline-specific KSAs and their cognate PLDs.

5 References & further reading

1 MacLean M & Miles C. (1999) **Swift action needed to close the skills gap in bioinformatics.** *Nature*, **401**, 10. **doi: 10.1038/43269**

2 Pevzner P & Shamir R. (2009) **Computing has changed biology – biology education must catch up.** *Science*, **325**, 541-542. **doi: 10.1126/science.1173876**

3 Abeln S *et al.* (2013) **Bioinformatics and Systems Biology: bridging the gap between heterogeneous student backgrounds.** *Brief. Bioinform.*, **14**, 589-598. **doi: 10.1093/bib/bbt023**

4 Rubinstein A & Chor B. (2014) **Computational Thinking in Life Science Education.** *PLoS Comput. Biol.*, **10**(11), e1003897. **doi: 10.1371/journal.pcbi.1003897**

5 Attwood TK *et al.* (2017) **Plug gap in essential bioinformatics skills**. *Nature*, **544**, 161. **doi: 10.1038/544161c**

6 Attwood TK *et al.* (2019) **A global perspective on evolving bioinformatics and data science training needs.** *Brief. Bioinform.*, **20**(2), 398-404. **doi: 10.1093/bib/bbx100**

7 Feldon DF *et al.* (2017) **Null effects of boot camps and short-format training for PhD students in life sciences.** *Proc. Natl. Acad. Sci.*, **114**, 9854-9858. **doi: 10.1073/ pnas.1705783114**

8 Welch L *et al.* (2014) **Bioinformatics curriculum guidelines: toward a definition of core competencies.** *PLoS Comput. Biol.*, **10**(3), e1003496. **doi: 10.1371/journal.pcbi. 1003496**

9 Welch L *et al.* (2016) **Applying, evaluating and refining bioinformatics core competencies (An update from the Curriculum Task Force of ISCB's Education Committee).** *PLoS Comput. Biol.*, **12**(5), e1004943. **doi: 10.1371/journal.pcbi. 1004943**

10 Mulder N *et al.* (2018) **The development and application of bioinformatics core competencies to improve bioinformatics training and education**. *PLoS Comput. Biol.*, **14**(2), e1005772. **doi: 10.1371/journal.pcbi.1005772**

11 Wilson Sayres MA *et al.* (2018) **Bioinformatics core competencies for undergraduate life sciences education.** *PLoS One*, **13**(6), e0196878. **doi: 10.1371/journal.pone. 0196878**

12 Via A *et al.* (2020) **Course design: considerations for trainers – a Professional Guide.** [version 1; not peer reviewed]. *F1000Research*, **9**, 1377 (document). **doi: 10.7490/ f1000research.1118395.1**

13 Tractenberg RE *et al.* (2020) **Guidelines for curriculum and course development in higher education and training.** *SocArXiv*. **doi: 10.31235/osf.io/7qeht**

14 Tractenberg RE. (2017) **The Mastery Rubric: a tool for curriculum development and evaluation in higher, graduate/post-graduate and professional education.** *SocArXiv*. **doi: 10.31235/ osf.io/qd2ae**

15 Fernandez N *et al.* (2012) **Varying conceptions of competence: an analysis of how health science educators define competence.** *Med. Educ.*, **46**, 357-365. **doi: 10.1111/j.1365-2923.2011.04183.x**

16 Bennett CJ & Walston SL. (2015) **Improving the use of competencies in public health education.** *Am. J. Public Health*, **105**, S65-S67. **doi: 10.2105/AJPH.2014.302329**

17 Miner KR *et al.* (2017) **The MACH Model: from competencies to instruction and performance of the public health workforce.** *Public Health Rep.*, **120**, 9-15. **doi: 10.1177/ 00333549051200s104**

18 Caverzagie KJ *et al.* (2017) **Overarching challenges to the implementation of competency-based medical education.** *Med. Teach.*, **39**, 588-593. **doi: 10.1080/ 0142159X.2017.1315075**

19 Tractenberg RE *et al.* (2019) **The Mastery Rubric for Bioinformatics: a tool to support design and evaluation of career-spanning education and training.** *PLoS ONE*, **14**(11), e0225256 **doi: 10.1371/journal.pone.0225256**

20 Via A *et al.* (2022) **Using the Mastery Rubric for Bioinformatics: a Professional Guide.** [version 1; not peer reviewed]. *F1000Research* (document). *In press.*

21 Bloom BS. (1956) ***Taxonomy of educational objectives: the classification of educational goals. Handbook 1: The cognitive domain.*** David McKay, New York, USA.

22 Anderson LW *et al.* (2001) ***A taxonomy for learning, teaching and assessing: a revision of Bloom's taxonomy of educational objectives***. Addison Wesley Longman, New York, USA. ISBN-10: 080131903X

23 Bologna Working Group on Qualification Frameworks. (2005) ***A framework for Qualifications of the European Higher Education Area.*** Ministry of Science, Technology & Innovation, Copenhagen, DK. ISBN: 87-91469-54-6

24 Magana AJ *et al.* (2014) **A survey of scholarly literature describing the field of bioinformatics education and bioinformatics educational research.** *CBE Life Sci. Educ.*, **13**(4), 607–603. **doi: 10.1187/cbe.13-10-0193**

25 Matser V. (2016) **BioExcel Deliverable 4.2 – Competency framework, mapping to current training & initial training plan.** *Zenodo*. **doi: 10.5281/zenodo.264231**

26 Jansen BJ *et al.* (2021) In ***Data-Driven Personas. Synthesis Lectures on Human-Centered Informatics.*** Carroll JM. (Ed) Morgan-Claypool, San Rafael, USA. pp1-10. ISBN: 9781636390680

27 Tractenberg RE. (2021) **Teaching and learning about teaching and learning: the Mastery Rubric for the Master Level (MR-ML).** *SocArXiv*. **doi: 10.31235/osf.io/md65v**

28 Ogilvie S. (2011) ***Institutions and European trade: Merchant Guilds, 1000-1800.*** Cambridge University Press, Cambridge, UK. ISBN: 978-1-139-50039-5

29 Prak MR (Ed.) *et al.* (2006) ***Craft Guilds in Early Modern Low Countries: Work, Power and Representation.*** Ashgate Publishing, Farnham, UK. ISBN: 978-0-7546-5339-4

30 Rashdall H. (1895) ***The Universities of Europe in the Middle Ages: Volume 1, Salerno, Bologna, Paris.*** Clarendon Press, Oxford, UK. ISBN-10: 1108018106

6 Acknowledgements & funding

GOBLET Professional Guides build on GOBLET's Critical Guide concept, using layout ideas from the Higher Apprenticeship specification for college-level students in England. The contents herein build on *The Mastery Rubric for Bioinformatics: a tool to support design and evaluation of career-spanning education and training* by Tractenberg *et al*., 2019[19].

7 Licensing & availability

The Guide is freely available for download via the GOBLET portal (**www.mygoblet.org**) and the F1000Research Bioinformatics Education and Training Collection (**f1000research.com/collections/ bioinformaticsedu?selectedDomain=documents**).

8 Disclaimer

Every effort has been made to ensure the accuracy of this Guide; GOBLET cannot be held responsible for any errors/omissions it may contain, and cannot accept liability arising from reliance placed on the information herein.

Organisations

GOBLET

GOBLET (Global Organisation for Bioinformatics Learning, Education & Training; **www.mygoblet.org**) was established in 2012 as a not-for-profit foundation to unite, inspire and equip bioinformatics trainers worldwide; its mission, to cultivate the global bioinformatics trainer community, set standards and provide high-quality resources to support learning, education and training.

GOBLET's ethos embraces:

- ***inclusivity****: welcoming all relevant organisations & people*
- ***sharing****: expertise, best practices, materials, resources*
- ***openness****: using Creative Commons Licences*
- ***innovation****: welcoming imaginative ideas & approaches*
- ***tolerance****: transcending national, political, cultural, social & disciplinary boundaries*

For general enquiries, contact **info@mygoblet.org**.

- Attwood *et al.* (2015) **GOBLET: the Global Organisation for Bioinformatics Learning, Education & Training.** *PLoS Comput. Biol.*, **11**(5), e1004281.
- Corpas *et al.* (2014) **The GOBLET training portal: a global repository of bioinformatics training materials, courses & trainers.** *Bioinformatics*, **31**(1), 140-142.

ELIXIR

ELIXIR is an intergovernmental organisation that brings together life-science resources (databases, software tools, training courses, cloud storage, *etc.*) from across Europe. The aim is to create a single infrastructure, making it easier for scientists to find and share data, exchange expertise, and agree on best practices: **elixir-europe.org**.

Through its Training Platform, ELIXIR is:

- providing services and tools for bioinformatics training, such as the Training e-Support System, TeSS (**tess.elixir-uk.org**), the ELIXIR Training Metrics Database (**training-metrics-dev.elixir-europe.org**) and the training Toolkit;
- supporting training providers across Europe by creating and delivering training for developers, researchers and trainers;
- building a sustainable training infrastructure.

Since 2015, the ELIXIR Training Platform and GOBLET have worked closely to promote and develop standards and best practices in bioinformatics training; the outcomes of this enterprise (peer-reviewed articles, training documents (Guides), posters, slides) are available from the F1000Research Bioinformatics Education & Training collection (**f1000research.com/collections/bioinformaticsedu**). Together, they have built a Train-the-Trainer (TtT) programme, which comprises a standard curriculum, associated training materials and well-trained instructors. To date, thousands of scientists have benefitted from this programme.

For general enquiries, contact **info@elixir-europe.org**.

About the authors

Allegra Via (orcid.org/0000-0002-3398-5462)[§]

Allegra Via has 20 years' experience in academia; her current focus is on educational psychology and curriculum development. She is a research scientist at the IBPM-CNR (Italy), the ELIXIR-IT Training Coordinator, and leads ELIXIR's TtT programme. She is Adjunct Professor at Sapienza University and University of Bologna, and expert consultant of Sapienza's Teaching Innovation and Quality working group, and of the University of Campania "Luigi Vanvitelli".

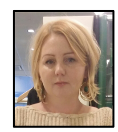
- Via A *et al.* (2017) **A new pan-European Train-the-Trainer programme for bioinformatics.** *Brief. Bioinform.*, bbx112.
- Via A *et al.* (2013) **Best practices in bioinformatics training for life scientists.** *Brief. Bioinform.*, **14**(5), 528-537.

Patricia M Palagi (orcid.org/0000-0001-9062-6303)[‡]

Patricia Palagi has more than 21 years' experience teaching bioinformatics, proteomics and programming in Swiss universities and international training programmes. Since 2013, she has led the SIB Training Group; she is the ELIXIR-CH Training Coordinator, leads ELIXIR's TtT programme, and is Instructor in ELIXIR's TtT courses.

- Palagi PM. (2020) **Lessons learned when moving SIB Training courses from in-person to online – anything in bioinformatics can be taught online, but not anyhow.** *F1000Research*, **9**, 1148.
- Garcia L *et al.* (2020) **Ten simple rules to make training materials FAIR.** *PLoS Comput. Biol.*, ***16****(5), e1007854.*

Jessica M Lindvall (orcid.org/0000-0002-5042-8481)[¥]



Jessica Lindvall is Associate Professor in Bioinformatics at Stockholm University, and joint head of facility for the SciLifeLab Bioinformatics platform in Sweden. She is the ELIXIR-SE Training Coordinator and an ELIXIR TtT instructor.

- Tractenberg RE *et al.* (2019) **The Mastery Rubric for Bioinformatics: a tool to support design and evaluation of career-spanning education and training.** *PLoS ONE,* ***14****(11), e0225256.*
- Gurwitz KT *et al.* (2020) **A framework to assess the quality and impact of bioinformatics training across ELIXIR.** *PLoS Comput. Biol.,* ***16****(7), e1007976.*

Teresa K Attwood (orcid.org/0000-0003-2409-4235)[†]

Terri Attwood is Professor emerita of Bioinformatics, with more than 25 years' experience teaching bioinformatics in BSc, MSc and PhD programmes, and in courses and summer schools, in the UK and abroad. She wrote the first bioinformatics text-book; her third book was published in 2016:

- Attwood TK & Parry-Smith DJ. (1999) ***Introduction to Bioinformatics.*** Prentice Hall. ISBN-10: 0-582-32788-1
- Higgs P & Attwood TK. (2005) ***Bioinformatics & Molecular Evolution.*** Wiley-Blackwell. ISBN: 9781405106832
- Attwood TK, Pettifer SR & Thorne D. (2016) ***Bioinformatics challenges at the interface of biology and computer science: Mind the Gap.*** Wiley-Blackwell. ISBN: 978-1-119-24345-8

Affiliations

[§]Sapienza University, Rome (IT); [‡]SIB Swiss Institute of Bioinformatics, Lausanne (CH); [¥]Stockholm University, Stockholm (SE); [†]The University of Manchester, Manchester (UK).