



Providing metabolomics education and training: pedagogy and considerations

Catherine L. Winder¹ · Michael Witting^{2,3} · Fidele Tugizimana^{4,5} · Warwick B. Dunn¹ · Stacey N. Reinke⁶ on behalf of the Metabolomics Society Education and Training Committee

Received: 23 June 2022 / Accepted: 11 November 2022

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

Abstract

Background Metabolomics is a highly multidisciplinary and non-standardised research field. Metabolomics researchers must possess and apply extensive cross-disciplinary content knowledge, subjective experience-based judgement, and the associated diverse skill sets. Accordingly, appropriate educational and training initiatives are important in developing this knowledge and skills base in the metabolomics community. For these initiatives to be successful, they must consider both pedagogical best practice and metabolomics-specific contextual challenges.

Aim of review The aim of this review is to provide consolidated pedagogical guidance for educators and trainers in metabolomics educational and training programmes.

Key scientific concepts of review In this review, we discuss the principles of pedagogical best practice as they relate to metabolomics. We then discuss the challenges and considerations in developing and delivering education and training in metabolomics. Finally, we present examples from our own teaching practice to illustrate how pedagogical best practice can be integrated into metabolomics education and training programmes.

Keywords Metabolomics · Education · Training · Pedagogy

✉ Stacey N. Reinke
stacey.n.reinke@ecu.edu.au

¹ Analytical & Clinical Metabolomics Group and Liverpool Training Centre for Metabolomics, Department of Biochemistry and Systems Biology, Institute of Systems, Molecular and Integrative Biology, University of Liverpool, Biosciences Building, Crown Street, Liverpool L69 7ZB, UK

² Metabolomics and Proteomics Core, Helmholtz Zentrum München - German Research Center for Environmental Health (GmbH), Neuherberg, Germany

³ Chair of Analytical Food Chemistry, TUM School of Life Sciences, Technical University of Munich, Freising-Weihenstephan, Germany

⁴ Department of Biochemistry, University of Johannesburg, Auckland Park, Johannesburg 2006, South Africa

⁵ International Research and Development Division, Omnia Group Ltd., Bryanston, Johannesburg 2021, South Africa

⁶ Centre for Integrative Metabolomics & Computational Biology, School of Science, Edith Cowan University, Joondalup 6027, Australia

1 Introduction

The term ‘metabolomics’ was first coined in 1998 (Oliver et al., 1998). Since then, substantial technological, methodological, and computational advancements have led to the maturation of the field (Kell & Oliver, 2016). Over the last 20 years, metabolomics has demonstrated utility in a wide range of contexts, including human health and disease (Beger et al., 2016; Rattray et al., 2018), nutrition (Brennan, 2017), environmental science (Bundy et al., 2008), plant and agriculture science (Fraser et al., 2020; Rai et al., 2019; Tinte et al., 2021), microbiology (Mashego et al., 2007), and synthetic biology (Hollywood et al., 2018). For a substantial period, the application of metabolomics was limited to academic researchers who had direct access to the required analytical and computational facilities. However, the arrival of analytical service providers has led to metabolomics becoming much more accessible to a wider range of applied life science researchers, further accelerating the growth of the field. Today, metabolomics is positioned as one of the key pillars in the systems biology (omics) sciences.

Despite the increased accessibility of metabolomics, two unique challenges remain for researchers. First, metabolomics is a highly multidisciplinary science (Fig. 1). To complete a metabolomics study from beginning to end requires expertise in the applied biological context, analytical and physical chemistry, quality assurance/control, cheminformatics, statistics and data science, bioinformatics, and biochemistry. Second, the metabolome itself is chemically highly diverse (in terms of physicochemical properties, concentration range and stability) and complex in that metabolites are highly interdependent and exhibit correlated dynamics. As a result, metabolomics analytical and data science workflows are difficult to standardize as often there is no single correct approach and each analysis is context specific. Accordingly, effective metabolomics researchers require a diverse set of knowledge in all areas of the workflow as well as specific skills in some parts of the workflow, not only cross-disciplinary objective knowledge but also a high level of subjective experience-based judgement.

The development of knowledge and skills can be supported through two non-mutually exclusive learning processes: training and education (Table 1). *Training* is a short-term process that facilitates the development of skills needed to perform a particular task (for example, how to use a particular software package). *Education* is a longer-term

learning process that focusses primarily on the development of theories and concepts and in some cases on their application. This type of learning typically occurs at schools, colleges, and universities. In an attempt to identify training needs in metabolomics, the Metabolomics Society and ELIXIR-UK jointly conducted and published a survey in 2015 (Weber et al., 2015). The key recommendations from this survey included: (i) development of face-to-face and online training courses in analytical metabolomics and bioinformatics, (ii) creation of funding opportunities to develop (inter-)national training networks, (iii) improvements in advertising and accessibility of training courses, and (iv) considerations for formal accreditation. Since 2015 the provision of metabolomics training has advanced considerably, by dedicated academics providing short course training and several centres specialising in the provision of both face-to-face and online training including the Liverpool Training Centre for Metabolomics, EMBL-EBI, Imperial International Phenome Training Centre, Birmingham Metabolomics Training Centre, the Metabolomics Workbench, the West Coast Metabolomics Centre, and Metabolomics South Africa (this is not a complete list of training centres available). These centres have the facilities and infrastructure via state-of-the-art instrumentation and computational infrastructure to provide a range of hands-on and online training

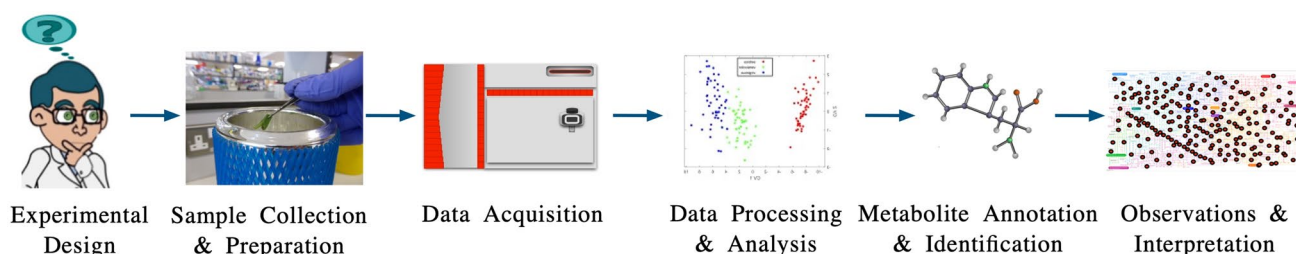


Fig. 1 Metabolomics workflow steps. The metabolomics workflow, from experimental design to biological interpretation, is shown. Different levels of (1) expertise and (2) awareness of high-level concepts are required for each step of the workflow

Table 1 A comparison of training and education

| Factor | Training | Education |
|----------------------|---|---|
| Provider | Primarily training centres and scientific companies which require no official registration (as would be required for education at a school or university) | Officially registered school, college, or university |
| Focus | Typically, hands-on training of skills to allow trainee to perform a task or set of tasks in their jobs | Provision of understanding concepts as primary focus with hands-on skill provision as a secondary focus |
| Course duration | Short-term (days) | Long-term (months to years) |
| Quality Assurance | Not required to operate | Normally required to operate |
| Students/trainees | Young adults and adults | Children, young adults, adults |
| Formal assessments | No | Yes |
| Informal assessments | Yes | Yes |
| Qualification | No (continuing professional development points possible) | Yes |

and fill training gaps highlighted in the survey. Geographical location of the trainer and trainees may limit accessibility of face-to-face training. Other training opportunities include conference workshops (for example, at Metabolomics Society conferences). While workshops provide an excellent opportunity to learn about focussed topics, such as new software developments, the time available and scale of the workshops may not provide sufficient opportunity to develop a higher-level knowledge and skill set.

One of the major challenges in delivering metabolomics training is to consider the varying levels of knowledge of participants. Researchers originate from varying backgrounds and although it may be expected that metabolomics teaching would be embedded in some undergraduate and post-graduate programmes (for example, biology degrees), this cannot be guaranteed. To the authors' knowledge, metabolomics teaching in education is emerging but remains limited; the authors are aware of only a small number of undergraduate and postgraduate programmes which include metabolomics in the curricula. Metabolomics-focussed papers has also recently started to appear in the educational literature (Boyce et al., 2019; La Frano et al., 2020; Sandusky, 2017). A search for metabolomics in postgraduate Master's programs across the globe also returned limited results; of 24,646 Master's programmes available globally and listed on the www.findamasters.com portal, only 50 programmes were listed using the search term metabolomics. Moreover, no single programme was focused on metabolomics only; instead, metabolomics was included as only one of multiple modules taught (FindAMasters, 2022). Indeed, the authors believe that a Master's course focused only on metabolomics is not appropriate because the volume of materials available is less than would be expected in a Master's course and that such a course would not be financially rewarding to the academic institutions as a low number of student registrations would be expected. In 2019 the Metabolomics Society Strategy Task Group published a survey of the membership which reported that 47% of respondents had been working in metabolomics for less than 5 years and 44% of respondents in trainee positions including undergraduate, Master's or PhD students and postdoctoral fellows (Zanetti et al., 2019). Owing to this relative inexperience of metabolomics researchers and lack of fundamental metabolomics teaching in the education setting, training and education remains a top priority for the field. The approach and mechanism of this training requires consideration to support all researchers working in the field.

Educating the scientific community in the application of metabolomics shares the same challenges as other disciplines such as bioinformatics, where potential users have different educational backgrounds and skills sets. The International Society of Computational Biology (ISCB) Educational Committee addressed the challenge by developing a

framework for training needs and curricula using a set of bioinformatics core competencies and mapping the competencies to typical user profiles (Mulder et al., 2018). The competencies and mapping to user profiles were refined via several workshops; however, the scoring approach was too ambiguous to allow meaningful classification until the competency levels were aligned to the pedagogical hierarchical cognitive model of Bloom's taxonomy (described in the next section). This illustrates the necessity of integrating pedagogical best practice when designing training and educational frameworks in contemporary research practice. To effectively support the development of knowledge and skills in metabolomics, pedagogical best practice should be implemented when designing and delivering education and training curricula; however, there is a lack of consolidated pedagogical guidance for the metabolomics community, including how to overcome metabolomics-specific challenges. In this manuscript we will describe key pedagogical principles as well as other practical considerations which we recommend should be applied in education and/or training. Pedagogical principles are reviewed from the education perspective and then adopted in the development of training courses. We also discuss additional considerations for developing and providing training courses focused on metabolomics workflows. Finally, we will provide case study examples of educational and training curricula from our own practice. As such, this manuscript will provide future and new teachers with a perspective on pedagogical and practical considerations, researchers with a deeper understanding of pedagogy as it relates to metabolomics, and experienced teachers with creative examples of integrating pedagogy with metabolomics education and training.

2 Principles of pedagogical best practice—constructive alignment

Before discussing considerations that are specific to the field of metabolomics, it is important to understand key foundational pedagogical principles. In this section we will discuss and provide examples of pedagogical principles applied in education, though many of the principles are equally important in training programmes. It is outside the goals of this manuscript to discuss in detail the general principles of pedagogy and we recommend the following resources to provide more detailed information (Anderson et al., 2001; Biggs, 1996; Hoque, 2017; Jarvis, 2012; Print, 1993).

2.1 Constructive alignment

A traditional view is that teachers often plan the content of their lecture(s), but students plan according to their exam requirements. This might lead to a misalignment of teaching

provided by the educator and learning required by the student. To avoid this, the concept of constructive alignment has been developed (Biggs, 1996). In this concept, learning outcomes, assessment, and teaching and learning activities are aligned. The three points to consider in constructive alignment are:

- Learning Outcomes—*What should students be able to do after the lecture/course?*
- Teaching and Learning Activities—*What content and methods shall be used to achieve the learning outcomes?*
- Assessment—*Will and how will achievement of the learning outcomes be assessed?*

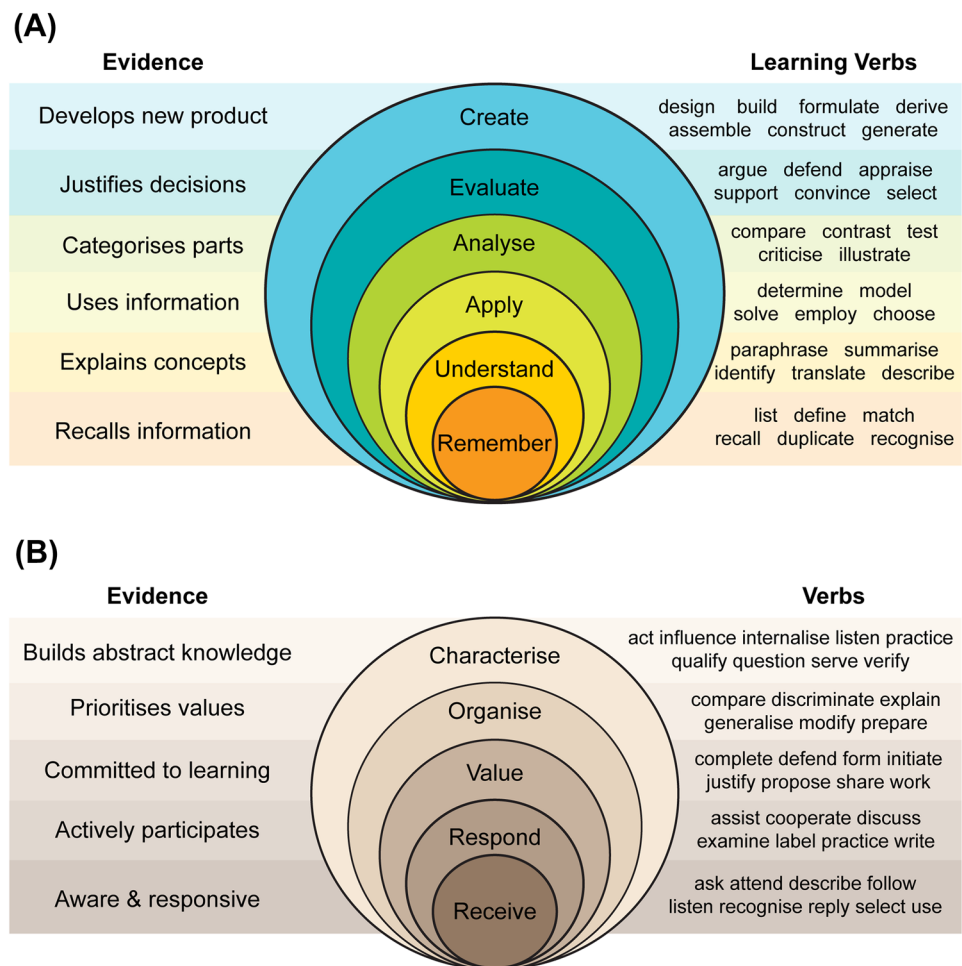
2.2 Bloom’s taxonomy—cognitive domain

It is important to define the competencies that students should have after successfully completing a teaching programme and to which level of expertise each competency can be acquired. A tool to order these levels is Bloom’s taxonomy, which comprises three different hierarchical models (cognitive, emotional, psychomotor). The cognitive domain

model [first reported in 1956 (Bloom, 1956) and revised in 2001 (Anderson et al., 2001)] consists of six different levels and is widely applied in curriculum development (Fig. 2). The model is hierarchical, with each level building off the last. Importantly, all learners, regardless of age or educational/training level, must start at the lowest level and progressively move upward as their learning evolves.

- The first three levels are often referred to as the lower-order thinking skills. The lowest cognitive level is ‘Remember’. This involves rote memory and is evidenced by activities such as recognition or recall of facts, definitions, and lists.
- The second level is ‘Understand’, which learners evidence by organising, summarising, generalising or describing concepts.
- The third level is ‘Apply’, where learners start to use their knowledge in a new way. Application is evidenced by modelling, solving, or choosing.
- At this stage, learners enter the higher-order thinking skills. In the fourth level, ‘Analyse’, learners can dissect

Fig. 2 Bloom’s taxonomy levels for **A** cognitive domain and **B** emotional domain. Evidence of each domain level is shown on the left, with indicative examples of learning verbs shown on the right. In both domains, learners must start at the bottom level and work through the respective levels to develop expertise in each level



concepts and ideas into parts and are able to compare them against other concepts and ideas.

- In the fifth level, ‘Evaluate’, decisions made can be justified. That is, learners defend their choices, and they can deliver arguments.
- The sixth and highest level in the taxonomy is ‘Create’. At this level, learners develop new ideas, concepts, or products. Since learners have evidenced all the lower levels, they can also defend the decisions they made in terms of design and conception. Without any doubt this is the level to which students should reach and it is the formal objective of many undergraduate programmes though is often only truly achieved for postgraduate PhD students.

2.3 Learning outcomes

In educational development, the first question that needs to be asked is: ‘What skills or knowledge should the learners be able to evidence upon successful completion of the teaching programme?’ Statements of this evidence are articulated through learning outcomes. To be most effective, learning outcome statements should integrate verbs that directly align to the Bloom’s taxonomy levels being targeted. By doing so, the learning can be planned to target an intended cognitive level. Learning outcome statements are most effective when they are clear and observable (Print, 1993). For example, learning verbs such as *explain*, *describe*, or *choose* clearly communicate expectations to learners, and completion of these outcomes can be measured. Contrarily, the verbs *know*, *understand*, and *demonstrate* are vague and cannot be directly observed.

In any educational programme, it is important to align the learning outcomes with the intended Bloom’s taxonomy levels. Learners must always progress through the cognitive levels, starting at the bottom. Thus, consideration must be made for what cognitive level is realistically achievable

based on prior competencies of the students and the time available for learning. For example, in a short course, it may be realistically achievable for a novice to learn how to perform Quality Assurance (QA) checks on a Liquid Chromatography–Mass Spectrometry (LC–MS) instrument, apply existing acquisition methods, or use existing data analysis tools. It may not be achievable for a novice learner to be able to troubleshoot LC–MS problems, create novel acquisition methods, or create bespoke data analysis methods in the same timeframe. Contrarily, these outcomes might be achievable for learners who have evidenced progression through the Bloom’s levels prior to the body of learning.

2.4 Teaching and learning activities

Once the intent of the learning (i.e. learning outcomes and cognitive levels) is defined, the next step is to consider *what* content will support the intent and evidence of learning as well as the sequence and mechanism of delivery. There are many ways to deliver content, which all fall under two overarching approaches (Jarvis, 2012): teacher-centred and student-centred learning (Fig. 3). Teacher-centred learning views students as ‘empty vessels’ and teachers as the source of knowledge. In this approach, teachers control the learning process and convey information to the students (for example, lectures and demonstrations). Teacher-centred approaches are effective for contexts where learning factually correct information such as rules, processes, and procedures is required. They are also appropriate in content-heavy curricula where a large volume of information needs to be delivered efficiently and often to many students. Student-centred learning approaches shift the focus away from the teacher and towards the student. Learning becomes less of a transmission of knowledge and more about an active process that the learner participates in (for example, workshops and project-based learning). This approach places much more

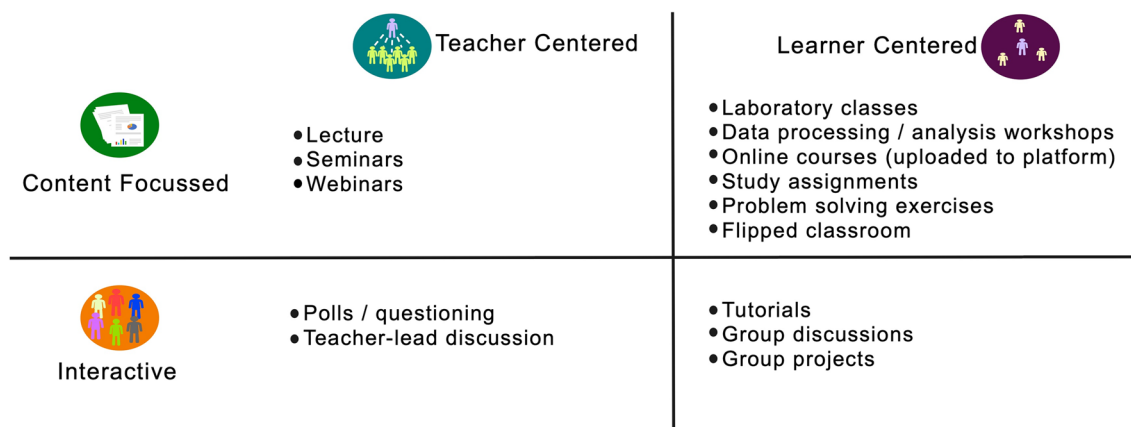


Fig. 3 Methods of delivering teaching and training material

accountability on the learner and it promotes autonomous learning. Typically, teacher-centred and student-centred approaches are used in combination to support both knowledge dissemination and active skills development. This is particularly relevant for metabolomics, where learners often come from diverse discipline backgrounds. An example of applying this dual approach is teaching metabolite annotation/identification. Students listen to an online pre-recorded lecture (teacher-centred) on the various physicochemical properties that support metabolite identification decisions. In class, students complete a student-centred activity called ‘Put Yourself on the Line’. The teacher draws an imaginary line along the length of the classroom. One end of the line represents ‘I am very confident in the identification of this metabolite’, the other is ‘I have no confidence in the identification of this metabolite’, and the rest of the line is a continuum between the two choices. Students are then presented an example metabolite peak with evidence to support its annotation/identification. They decide how confident they are in the identification and stand at the corresponding place along the line. Once all students are standing on the line, they are invited to articulate their decision for choosing their position. This process may be repeated with several examples to challenge students and develop higher level ‘analyse’ and ‘evaluate’ skills.

A range of teaching and learning tools can be applied to facilitate content delivery and supplement more traditional teaching tools and also maintain interest, accommodate different learning styles, and improve the learning experience. Specific examples that are relevant to metabolomics include:

- Applying live polls during teaching/training sessions to allow (1) students to assess their level of knowledge and compare this to their peers and (2) teachers to determine the current level of learning achieved. For example, a live poll may ask students to report on the criteria they would use for data cleaning. The results would provide different perspectives on what the learners find important (example: peak shape, quality metrics such as relative standard deviation, a confident identification).
- Providing online lectures and/or recordings of classroom sessions allows students to watch in their own time and at their own speed with the ability to rewind and fast forward. The typical 50-min classroom lecture can be divided in to smaller and multiple lectures with linking text which can improve students’ interest and learning. This can be particularly useful for a metabolomics audience, where learners have different levels of prior knowledge. Recordings allow novices to spend more time on learning the content.
- Providing links to audio-visual (AV) resources to supplement teaching. For example, providing a link to a YouTube video focused on the learning objective for which

a lecture was provided by the teacher allows the same concept to be described by two different teachers and most probably in two different ways. If the first teacher’s approach was not best suited to the learning style of the student, then the YouTube approach may be successful. Again, this can be particularly useful for novice learners.

- Computational notebook-style live scripts (example: Jupyter Notebooks, R markdown, MATLAB live scripts) can help students develop their coding capabilities. With live scripts, teachers can provide background on the algorithms and functions (including links to external resources) as well as suggestions for changing input arguments to promote experiential learning. Mendez et al (2019) provide a tutorial review on using Jupyter Notebooks to support metabolomics.
- Online tools and resources for hands-on teaching. For example, web-based software requiring no downloading and installation, such as MetaboAnalyst, and which can easily be applied by one or one hundred students. Microsoft Azure Labs (<https://labs.azure.com/>) offers a virtual machine (VM) solution, where teachers create a master VM with all required software and files loaded. The template is then published and a private VM clone is created for each student. This enables students to access the required computational resources at any time and from anywhere in the world. This is particularly relevant for teaching metabolomics in contemporary contexts. Metabolomics requires access to and installation of specialised software, and large volumes of data, which is practically difficult to support in remote teaching contexts. For an example that illustrates how Microsoft Azure Labs has been used in teaching, see the ‘Case Studies’ section below.
- Open access data. Metabolomics data collected applying NMR spectroscopy and LC–MS is now available in metabolomics data repositories [e.g. MetaboLights (Haug et al., 2012), Metabolomics Workbench (Sud et al., 2015) and GNPS (Wang et al., 2016)] and can be used for hands-on teaching. The reanalysis of deposited data will be applied more frequently in future research and so teaching the next generation in its use is important.

2.5 Assessment

Assessment is a critical component of the learning process in educational settings. It not only evidences achievement of learning outcomes (Biggs, 1996) but it is a critical component in supporting student learning (Jarvis, 2012). There are three main reasons why assessment is included in educational curricula.

1. Assessment *for* learning. This provides the teacher diagnostic information which they can use to inform, improve, and direct their teaching.
2. Assessment *of* learning, which evidences learners' achievements of learning outcomes.
3. Assessment *as* learning or self-assessment. This provides students an opportunity to not only monitor their own progress but reflect on their learning to direct future activities.

There are two primary types of assessment: formative and summative. Formative assessment occurs before and during the learning process and provides students with feedback to improve their learning during the learning cycle. Formative assessments are typically low stakes (or not marked) and include activities such as in-class or online discussions, small and frequent quizzes, and short writing assignments. Summative assessments occur at a later stage or the end of the teaching programme and are used as evidence that learning outcome have been achieved and to award achievement. These assessments are typically high stakes and include assessments such as exams, project reports, literature reviews and portfolios. There is also a third type of assessment, self-assessment. A range of tools are available to embed self-assessment into teaching and training to facilitate self-assessment and self-direction of learning. The range of assessments are shown in Table 2 along with a classification of the assessment type, and examples when typically used. Specific examples of assessments used in metabolomics curricula are described in the Case Studies section below.

3 Principles of pedagogical best practice—other considerations

3.1 Understanding prior knowledge

“To teach is to learn twice” (Joseph Joubert)—A prominent fact to consider when teaching is the “expert blind spot”. Metabolomics teachers are often experts who have years or even decades of experience in their field. It can be easy to take this knowledge and experience for granted and forget how difficult it is to learn something new. As metabolomics is a multidisciplinary field involving aspects from biochemistry, analytical chemistry and bioinformatics, heterogeneity of the students' knowledge and experiences is an important factor. It is important to not only factor this in during planning but also to regularly check with students that the pace and depth is suitable. Providing additional resources (for example: readings, online resources) can help students to close the gap and consolidate their knowledge away from teaching/training sessions.

3.2 Bloom's taxonomy—emotional domain

A more recent development of Bloom's taxonomy has considered the emotional levels of the students and their ability to learn, specifically that the student must be emotionally involved and motivated to participate in the learning process (Krathwohl et al., 1964). The teaching and learning activities and tools can be developed to facilitate the emotional involvement and motivation of the student. There are five levels and as with the cognitive levels, each new level builds on the previous levels (Fig. 2).

- Receiving: the student passively approaches the teaching process by paying attention to the teaching process. This level is about the student's memory and recognition of materials.
- Responding: the student actively participates in the learning process through reaction to the teaching stimulus.
- Valuing: the student understands the value of learning to increase their knowledge.
- Organizing: the student can integrate different information, ideas, and values to allow comparison and elaboration of what has been learnt.
- Characterising: the student can build abstract knowledge.

4 The challenges and considerations in education and training development and operation

Pedagogical best practice must be considered within the constraints of the teaching context. In this section we will discuss challenges and considerations that are relevant to the metabolomics context. Although both education and training will be covered, we will focus more on training course development and operation. Training courses form an important part of Continuing Professional Development (CPD). As metabolomics is a relatively new discipline, most of the formalised metabolomics education/training to date will have been performed via post-graduate education or CPD-based training. Undertaking CPD within your role ensures your knowledge stays relevant and current. Short courses are used as an intermittent booster to fast-track knowledge in a specialised subject but also provide access to experts. With new approaches such as metabolomics, CPD-based training courses provide a mechanism to teach new knowledge and skills to those who completed their formal education many years ago. Many of the concepts discussed above for education can also be related to training, for example, Bloom's taxonomy and teaching tools and the teaching approaches summarised in Fig. 3. However, the goals of education and training are different. When developing a training course, the content will typically include a narrower syllabus than

Table 2 Types of assessments that are routinely used across the teaching and training environments in metabolomics

| Assessment/activity | Classification* | | | | Description | Examples |
|---|-----------------|----|---|----|--|--|
| | SA | FA | S | TR | | |
| Examination | x | | | | Formal procedure may employ MCQs, short answer and essay questions. Applied to award qualifications | Final year examinations in undergraduate teaching |
| Test | x | x | x | | Tool to measure knowledge levels and adjust teaching material/delivery accordingly. Various formats may be employed including MCQs, short and long answer questions | Mid or end of module assessment in undergraduate and post-graduate teaching |
| Thesis/Dissertation | x | | | | An extended written document focussed on a specific subject, demonstrates research and contribution of new knowledge, theories or practice to a specific field | BSc or MSc thesis PhD dissertation |
| Viva/Oral Defence | x | x | | | Oral examination a body of work in front of a panel of experts and chair. Frequently used in combination with a written dissertation or report as part of a final assessment or to assess progress at specific milestones | A PhD viva |
| Project report | x | | | | A structured report to present the strategy, results, and conclusions of a project, may include a laboratory, computational or problem-solving project | Laboratory project |
| Portfolio | x | x | | | Collection of student work completed over time. Learning is structured such that students complete various tasks throughout the course. Evidence of task achievement is then collated into a portfolio, may include a reflective element | Undergraduate/masters portfolio to document completion of modules in data processing, analysis and metabolite identification |
| Reflective assignment | x | x | | | Assignment that guides students to consider, articulate their experience, and inform future practice. May be a stand-alone assignment or part of another assignment (such as a portfolio) | Undergraduate assignment to define and consider how a work-based placement has improved their skill-set for future employment |
| Presentation (oral, poster, elevator pitch) | x | x | | | Can be used in a variety of different ways. Facilitates development of oral and visual communication skills | Oral presentation at end of project to present project findings, often applied in undergraduate and post-graduate taught courses |
| Group task | x | x | | | A collaborative activity where students work in groups to perform tasks or solve problems, may involve a short task or project over several days or weeks | Problem-based activity to design a metabolomics experimental plan |
| Laboratory/skills assessment | x | x | | x | A laboratory-based skills assessment to test the practical skills and competence level of students. The assessment will be based on a set of transparent criteria outlined in advance | Undergraduate laboratory practical exam in range disciplines including biology, chemistry, biochemistry |
| Quiz | x | x | x | x | Low-stakes assessment. Options include short answers or input the missing word (cloze question). If implemented using an online platform, auto-feedback and auto-grading can be applied, providing students immediate feedback upon completion (or deadline) | Cloze questions to add in missing word as part of an online quiz |

Table 2 (continued)

| Assessment/activity | Classification* | | | | Description | Examples |
|--|-----------------|----|---|----|---|---|
| | SA | FA | S | TR | | |
| Multi-choice question (MCQ) quiz or test | x | x | x | x | Tests knowledge and understanding, applied on online platform to support self-assessment in education and training sections. Implement auto feed-back to direct students back to specific areas of course. Question design important to test students' knowledge, not providing obviously incorrect answers | MCQ's embedded in online course employs auto-marking and auto-feedback |
| Discussion activity | | | x | x | Face-to-face or asynchronous online activities. Questions are posed by the facilitator. Can be used for a variety of reasons, including to promote inclusion between students, extension of learning, development of critical thinking skills, and compilation of resources | The use of discussion boards in data analysis courses to explore and evaluate the use of data analysis approaches across the field |
| Polls | | | x | x | Live or asynchronous survey. Typically, multiple choice or true/false answer options. Results can be conveyed back to learners | Live poll during lectures to stimulate thought and discussion, a good option when considering quality assurance and quality control in metabolomics. Facilitates interactive learning in lectures |
| Directed exercise | | | x | x | Data processing activity | Workshops on how to code or data analysis tutorial using web-based software such as MetaboAnalyst |

*Classification of assessment abbreviations: SA summative assessment, FA formative assessment, S self-assessment, TR Tutor review to direct learning

in the education course to impart specialised knowledge, skills, and direct students to resources. When developing such courses considering the audience and challenges specific to the subject area and considering what cognitive level is realistically achievable within the course are very valuable and are discussed below.

4.1 Challenges and considerations for education and training

4.1.1 All omics are not the same

Many different omics research areas and tools are available for 'discovery-based' biological research including genomics, transcriptomics, proteomics, and metabolomics. The general research goal of these omics techniques is to identify biochemicals which are involved in biological mechanisms or have the potential to act as biomarkers. Importantly, and often forgotten, is that different tools are applied in the research laboratory and computational office for different omics disciplines. For example, the biological roles of metabolites and speed of metabolism can be different to other types of biochemicals, therefore sampling and storage methods can be different for metabolomics compared to other omics techniques and so an expert in say proteomics is not an expert in metabolomics (and vice versa). This is an example of how users require specialised knowledge in for example, sample collection, storage, and preparation. Scientists who transition from another 'omics disciplines will have transferable skills, such as the ability to operate a mass spectrometer or experience of working with large data sets but will require metabolomics-specific training to understand the challenges of studying the metabolome and develop an appreciation of how the multiple omics approaches differ. For new users to the field, who are not working in an established metabolomics research group and therefore cannot learn from their peers, hands-on laboratory, bioinformatics, and data analysis training is an invaluable part of their professional development.

4.1.2 Research involving metabolomics is multidisciplinary

A wide range of different scientific fields can be involved in a study which applies metabolomics research tools including biologists, clinicians, plant biologists, analytical chemists, physicists, bioinformaticians and computational biologists. Therefore, one unique challenge to consider when developing training is the diverse group of scientists working in or applying metabolomics in their research, understanding the varying background knowledge and skills. Whether a scientist is a specialist in one part of the workflow or is involved in the entire pipeline, they require an appreciation of all steps in the metabolomics workflow. For example, when a statistician is

reviewing the outputs of the data analysis it is useful to have an appreciation of the experimental pipeline and steps which may introduce sources of variation. A clinician will often collaborate with a core metabolomics group to perform the experimental and data analysis work. So, whilst an in-depth understanding of some skills (for example, operating a mass spectrometer) is not required, a good background knowledge of the subject, experimental design and an understanding of the challenges is essential. Accordingly, consideration must be made about whether one trainer/educator can reasonably cover all aspects of the course content. The authors recommend that multiple trainers with different expertise areas are involved in a single training course and that trainers limit their teaching scope to content that they have deep expertise in.

Educating and training scientists with no or limited expertise requires specialised courses that:

- Consider and accommodate different scientific backgrounds and levels of knowledge
- Address the needs of the trainee at different stages of their careers
- Provide training that is relevant to the needs of the scientist
- Supports transitioning between areas of specialisation, for example, an experimental scientist moving from a biological focus to analytical chemistry or learning to develop computational tools.

We provide two examples in the case study section to demonstrate how to overcome the diverse backgrounds of students in (1) a teaching-focussed postgraduate course and (2) supporting research-based students with minimal knowledge in metabolomics.

4.1.3 Feasibility

Learning is a complex process. It can be easy to overestimate what can be achieved both by the teacher and the learners. Two primary factors that are often underestimated by inexperienced teaching staff are time and cognitive load. Effective scaffolding and development of learning activities takes a substantial amount of time. It is important to consider what can realistically be achieved in development from a workload standpoint. Every iteration of delivery will identify new areas of improvement (through trainee feedback and trainer reflection), which can be developed incrementally. It is also easy to underestimate the amount of time it will take trainees to perform the learning activities, and this should be accounted for in the planning stage. Cognitive load is also important to consider. According to cognitive load theory (Sweller, 1988), there is a limited capacity of the working memory at any given time. If the working memory is overloaded with too many simultaneous tasks, learning is

impaired. This is important for metabolomics, as students are often required to learn both concepts and applications. For example, students learning about statistics must learn both statistical concepts and how to calculate statistics using software. Learning to use statistical software with a graphical user interface (GUI) such as SPSS is less cognitively demanding (at least for beginners) than using a command-line interface (CLI) software such as R. Therefore, efforts to reduce cognitive load will allow learners to focus on the key learning objectives. The use of MetaboAnalyst in the delivery of a data analysis course is an example where participants can focus on the reasons and correct use of the processing steps without having to learn how to programme.

4.1.4 Is the infrastructure available?

Anyone can theoretically operate a training course. However, appropriate infrastructure is required to deliver high quality training and learning. For online courses a training platform with the capabilities to present training material delivery in different formats and facilitate interaction between trainees and trainers; for example, providing only a series of YouTube videos with no interaction will not facilitate engagement and higher-level learning (as defined by Bloom) in all participants. The delivery of face-to-face courses will require a training room and possibly a laboratory or computational cluster. The training environment has to be appropriate (e.g. comfortable, appropriate temperature, no loud noises) with access to state-of-the-art instrumentation and software. Many users would like to gain hands-on sample preparation and data acquisition training. Metabolomics training is frequently provided by experts working in the field however, the ability to operate hands-on laboratory courses is somewhat limited and this type of metabolomics training are limited to a small number of centres and organisations; examples include the Imperial International Phenome Training Centre, Birmingham Metabolomics Training Centre, UC Davis West Coast Metabolomics Center, the Liverpool Training Centre for Metabolomics and Metabolomics South Africa (MSA). Providing access to state-of-the-art instrumentation is a challenge. Instruments in research laboratories are extensively used to generate data for multiple research projects and so are not always readily available to use in training courses. In addition, highly trained operators who have other commitments are required to set-up and run the hands-on training. The provision of infrastructure to provide lecture-based or computer workshops is easier to facilitate than laboratory space. The ideal scenario is to mirror the real-life metabolomics workflow and deliver face-to-face training via a combination of lectures, lab classes and workshops allowing attendees to gain practical experience. Providing dedicated time for questions and discussion helps the trainee to reinforce learning objectives and receive

advice and feedback on specific problems. In the author's experience, having time for trainees to discuss their research with the trainers so to provide advice and support is essential in any training course.

4.1.5 Course development and revision

Metabolomics, as with the other 'omics is a rapidly developing field. A finding of the survey by Weber et al. (2015) reported that researchers wish to receive training in key areas of scientific development within the field and would like hands-on experience. Developing and updating courses that include hands-on training, laboratory or data workshops will require a greater time investment. In addition to methodology and skills-based activities, CPD-based training provides a platform to communicate and foster an understanding of key issues in the field and share experience from more experienced users. For example, the tools and techniques applied in metabolite annotation/identification have undergone rapid development over the last few years. Trainees benefit from understanding how to use the new approaches but to also understand and communicate the limitations and confidence in metabolite annotation/identification to the wider audience. Other key issues to share with new users are the importance of open data sharing and data repositories such as the ELIXIR resource *MetaboLights* (Haug et al., 2012) or *Metabolomics Workbench* (Sud et al., 2015), and an understanding of the minimum reporting standards conceived by the *Metabolomics Standards Initiative* (MSI) (Fiehn et al., 2007). The adoption of open data standards assists scientist to better share and re-use data. The FAIR (Findable, Accessible, Interoperable, Re-useable) guiding principles for data management and stewardship were developed to support both data producer and publisher. Development of metabolomics data repositories has provided the community with a new resource to reuse data in the development of bioinformatics and data analysis tools. Good research data management enhances the reusability and data integration and is an important area of training within the metabolomics community. Within the UK the EMBL-EBI metabolomics team promote the importance of open data sharing and data repositories through numerous training activities. The global success of these repositories will only be maintained by continued support across the metabolomics community to promote open science and data sharing. Encouraging existing and new users to dedicate time in submitting data using guiding principles such as FAIR and educating users in the tools developed to facilitate the process.

4.1.6 Course impact and evaluation

Although most training courses do not contribute towards a formal qualification it is desirable for attendees to receive

a certificate of attendance or completion to record course attendance and is often required by the attendee to document their professional development and justify course fees. A European Credit Transfer System (ECTS) operates for students from the European Higher Education Area whereby they can gain credits towards qualification from attending short training courses. Accreditation of training courses by professional bodies provides official recognition that a course reaches a particular standard. It is a mechanism to demonstrate the value of the courses and is useful in course marketing. The Royal Society of Chemistry, Royal Society of Biology and Royal Society of Medicine in the UK are examples of professional bodies that accredit courses within the remit of metabolomics. For example, courses provided by the Imperial International Phenome Training Centre are accredited by the Royal Society of Medicine. The final consideration when performing any type of training is how to evaluate the success. Criteria may include the level of participation and popularity. Feedback surveys are often used but are provided at the end of the course, so feedback is usually focussed on course logistics rather than the longer-term impact of the training. Trainers can also inform training development and revision through their own observations, sometimes enabling real-time adjustments to be made. Finally, if formal assessments are included in the programme, they can also offer insight on gaps between curriculum intent and achievement against the learning outcomes. While deciding on adjustments to make, it is important to consider how the adjustments will enable achievements of the learning outcomes and what is practical within the context of time and resources.

4.2 Challenges and considerations specific to training courses

4.2.1 Trainees have different levels of metabolomics expertise

When undertaking any type of teaching, training or even a scientific presentation you must know your audience and consider the limitations of what can be achieved. To address the training needs of the scientific research community a range of courses from introductory, through intermediate to advanced levels are required. As discussed in the *Principles of Pedagogical Best Practice* section the starting point of any course development is the learning objectives. Introductory courses start from the beginning and have the objective to define and explain concepts to scientists who have not applied or have minimal exposure to metabolomics in their research. For example, a one-day lecture-based course which defines the basics of when metabolomics should and shouldn't be applied, what metabolomics is, the experimental workflow, and appropriate considerations for the trainees

to develop a core understanding of the topic is a good starting point. It is appropriate for those who are interested in applying metabolomics either in their own research groups or in core metabolomics groups and applies the two lower levels of Bloom's cognitive domain (Fig. 2), 'remember' and 'understand'. More specialised courses can be operated at intermediate to advanced levels addressing specific areas such as mass spectrometer maintenance and operation, data analysis or metabolite identification, and vocational based training to cater for specific groups of trainees generally interested in applying metabolomics in their own research groups. These courses look to develop learning at all levels of Bloom's cognitive domain including 'create'. Multiple learning opportunities are required to develop a complete understanding of the approach, be it via self-centred learning, learning from experts in 1:1 settings or more advanced CPD-based training. With practical components such as instrument operation, or data analysis hands-on testing of the approaches outside of the teaching environment in-between formalised training are essential to maximise the learning experience during the training courses and develop a deep understanding of the subject. In our (CW and WD) experience of operating metabolomics-based training courses since 2015 attendees benefit from the opportunity to ask questions months to years after attending the course. Vocational training is targeted towards the requirements of a specific group. For example, both clinical and environmental scientists benefit from dedicated sessions to address particular challenges they encounter when collecting their different types of samples in different environments (clinic, bench, remote field site). They may share the overarching challenge of how to quench or cool and transport samples during collection, but specific conditions (for example a clinic or a remote field site) will vary and influence procedures that may be applied. Training by the community is content focused and is typically delivered by a small number of centres or groups specialising in the field, at conference workshops or by instrument manufacturers. This specialised training can also be supplemented by other skills training to allow a researcher to develop the diverse range of skills required. For example, skills training such as software carpentry courses in the programming languages R or Python are a great resource and a valuable part of the training portfolio.

4.2.2 The number of trainees registered on a course

For efficiency and to minimise the number of training courses needed to be operated, each training course ideally would be available to many trainees. However, there are several reasons that influence the maximum number of trainees on a particular course. In laboratory-based training, where a demonstration of a technique or instrument use is required then having many people trying to view this process will

result in some trainees not being able to view fully and learn appropriately. The authors (CW and WD) have found that a maximum number of four trainees per instrument is appropriate when viewing the instrument operating computer to allow each trainee good vision and the time to practice the process during the training session. This may vary for different training environments. Applying such limits will reduce the number of trainees than can attend each course; however, multiple different processes can be operated at the same time with different trainees working on one process and then moving on to the next process. In the authors experience (CW and WD) running face-to-face courses with a maximum of 8 trainees works well in relation to high-level visual training but also in group dynamics and interactions. For online courses a larger number of trainees can attend, and the authors (CW and WD) typically encourage more than 50 trainees for a scientific online training course. With online courses such as Massive Open Online Courses (MOOC), a key part of the learning experience is through the discussion forums and peer learning, it is therefore beneficial for large cohorts to enrol on the course to facilitate discussions. A high number of trainees may not be beneficial in all circumstances for example, if a live component is included in the training, then appropriate time or training support will be required to ensure all questions are answered during the session. Even with courses that are not operated in real time sufficient resources are required to answer questions and facilitate large cohort courses, this can include PhD students working in the field to work alongside the more experienced trainers.

4.2.3 Should education and training be provided face-to-face or online?

Whether to operate a course face-to-face or online is an important consideration and may be influenced by the type of training to be delivered. Different online approaches can be applied in the learning environment including webinars, distance learning as either live or recorded videos and MOOC's that balance good content with context. The use of online approaches in the training environment has increased in popularity from 2011 with the widespread development of MOOC's. Delivering online teaching rapidly increased in 2020 when due to necessity online distance learning was employed for most undergraduate and post-graduate courses. The rapid development of online tools to attend meetings, learn and network in addition to social media provides new mechanisms that have been embraced by the scientific community to develop global networks and can be applied to enhance the online training experience.

Online training can reach a larger audience and is a useful format although not appropriate for all steps in the metabolomics pipeline. The face-to-face format is more appropriate

for laboratory-based activities where specialised facilities and equipment are required and results in a high level of interaction between trainers and trainees. Trainees often state that one of the reasons for attending training courses is to develop professional networks. Live discussions, question and answer panels and social events in face-to-face courses allows trainees to network with experts and build networks. In addition to laboratory processes, training in bioinformatics tools is frequently performed during face-to-face training but could also be provided via online videos with the potential to reach more trainees. With this approach, interaction and discussion must be facilitated through other mechanisms. In live online sessions, interaction can be synchronous with verbal discussions taking place at the whole group level or in virtual breakout rooms. Asynchronous online courses do not provide synchronous interaction observed in face-to-face courses. Alternative approaches can be employed to ask and answer questions, for example, the training material can be hosted on a suitable platform with discussion boards, comment sections and tasks to facilitate discussion between trainees.

Developing online courses require a greater time investment than producing face-to-face material. It is estimated that for every hour of training it takes 6 h to develop face-to-face training and 15 h to develop online training. However, online courses are accessible to a wider global audience, as attendees do not need to travel, and many online courses are uploaded to a learning platform so the user can complete the course in their own time frame. The intensity of the training can vary between the face-to-face and online format. Face-to-face training courses and some short online courses are operated over a small number of hours (for example, a scientific webinar) or over a small number of days with training for several hours each day. Online training courses allow for the training materials to be provided over a longer period (multiple weeks) with fewer hours per week required by the trainees and so can be accommodated into the routine working pattern rather than taking several days away from work.

The authors apply both synchronous and asynchronous in their course delivery and find blended approaches using online and face-to-face formats useful. Flip-style teaching where participants can view recorded videos of introductory material and concepts in advance of attending a tutorial or short face-to-face training course is useful when training participants from diverse scientific and knowledge backgrounds. A range of training materials can be provided in advance to allow all attendees to start for a common point of understanding. The face-to-face time can then be spent on learning outcomes at the higher level of Blooms cognitive domains. We provide two examples in the case study section to demonstrate the application of online training. One example demonstrates a blended course delivery employing both face-to-face and distance learning on a Master's courses

and the second example uses an online platform to deliver asynchronous training.

4.2.4 The costs of developing and operating training courses

Training courses, like any educational process, requires (dedicated) resources to develop and operate. These resources include one or more of the following: trainer time/salary, training room and laboratory space, laboratory consumables and scientific instrumentation, training materials (written or online), software, PCs/laptops, access to data servers and catering requirements. High quality training courses require trainers that are not only experts in metabolomics but also experts in course design and provision. The environment should be conducive to learning so with minimal background noise and availability of state-of-the-art instrumentation and software. A cost recovery model is often the route to support training centres with a fee charged to attend courses, this may be offset for the trainee when bursaries are available from research council funding. Some courses can be operated for free, typically short courses operating up to a single day. Registration fees can be paid by attendees and is a common approach for laboratory-based or online courses. However, this can be limiting to early-career scientists for example where the finances are not available. A third option is for the trainers to apply to research councils for bursaries which can be applied to trainees who meet specific criteria. Costs of a training course are dependent on the resources required and length of course; face-to-face courses generally have higher registration fees compared to online courses which require fewer resources though costs related to the training platform applied must be considered.

5 Case studies

In this section, the authors provide examples from our own training and educational practice to illustrate how pedagogical best practice can be applied to metabolomics whilst considering practical challenges.

5.1 Developing face-to-face taught postgraduate university course to introduce the metabolomics workflow to students from diverse educational backgrounds

The School of Life Sciences at the Technical University Munich (Author MW) offers an elective course on metabolomics in the international Master of Science, Nutrition and Biomedicine course. Since metabolomics is only partially covered in other course components, this elective course aims to close this gap for interested students. This course

covers all elements of the metabolomics workflow from study design, analytical technologies and data analysis and interpretation and aims at higher cognitive domains according to Bloom's taxonomy. On completing the course students should be able to *compare* analytical methods used in metabolomics and *evaluate* the advantages and limitations of each method to solve a specific scientific question. Students *design* their own metabolomics experimental plans and *evaluate* their peers' experimental plans. In order to enable students to develop the required skills this course is split into a one-hour lecture and two hours project work session each week.

In developing the course contextual challenges were considered due to the high heterogeneity in student background (international students from different countries with different bachelor studies and degrees). The diversity in the background of the students is taken as an advantage for this course. The entire course work is based on the concept of problem-based/problem-oriented learning. This type of learning is well suited for higher cognitive domains and allows students to develop independent and critical thinking. The course is structured into 3 different sections and different aspects of metabolomics are always discussed in the frame of a specific problem. Each section covers about 4–5 weeks. In the first section the lecturer uses the first problem to introduce basic concepts of metabolomics mostly on its own with partial input from the students. After this introduction, the second section is used to allow students to study a common problem and start to develop their own ideas for metabolomics workflows with partial input from the lecturer. Students work on different aspects in groups and present their results and solutions to the entire audience for open discussion. In the last section, groups of 2–3 students work on individual topics given by the lecturer. These topics also serve as an assessment for the students. It was not possible to offer a laboratory component in the course, this leads to difficulties in explaining specific parts of the workflow. At least laboratory visits of a running metabolomics laboratory are conducted to get an impression on the laboratory work. However, selected interested student are offered an internship or a master thesis position afterwards. Presentations and discussions throughout the lectures enable students to develop presentation skills, which is required for the assessment as well in future settings, e.g. during defence of a Master's thesis or other project work. The assessment of this module is conducted in the form of project group work. It consists of an oral presentation of 5–7 min per person and submission of a maximum 6-page long research paper. The group selects from different possible topics introduced by the lecturer to the audience. The research paper is a method to measure the overall understanding of the stated problem and their ability to solve complex problems, analyze the current state-of-art and develop novel solutions. The oral

presentation allows students to present their results to a wider auditory and subsequent discussion is a mean to measure their understanding of the scientific subject. Through the selected modes of assessment students can develop important skills such as literature search, extraction, and condensation of information.

5.2 Developing education and training to support metabolomics research-based programmes with minimal education in the metabolomics discipline

Metabolomics, as a multidisciplinary science, is not yet featured in the curricula of high educational institutions and universities in South Africa. Metabolomics research, in South African academia, is conducted at and through the Master's and PhD levels. Thus, the education and training in metabolomics are part of these research-based programmes. The candidate (a postgraduate student) and his/her supervisor (PI or mentor) share responsibilities. To illustrate this, the case of metabolomics (part of the research performed in the Department of Biochemistry) at the University of Johannesburg (Author FT) is highlighted.

Modules such as advanced analytical techniques, taught at the BSc Honours level (a 1-year postgraduate degree, before a Master's course) introduces students to a range of analytical techniques some of which are used in metabolomics studies, e.g. chromatographic techniques (LC and gas chromatography), mass spectrometry, NMR and other spectroscopic techniques. This module on advanced analytical techniques aims at cognitive domains of Bloom's levels from understanding to evaluate. After the course, the student should demonstrate an advanced understanding of principles of different analytical techniques, and their applications and limitations. The course runs throughout the first semester and uses both teacher-centred and student-centred approaches. The module is designed to have a 2-h session each week of a lecture (theory and demonstration) and hands-on training on instruments (e.g., liquid/gas chromatography-mass spectrometry). Assessments are performed throughout the course via assignments, tests, exam (with theory and application questions) and a competence-based report on hands-on training tasks. The first year of the master's (or/and PhD) programme is used to train and teach new concepts and aspects of the metabolomics workflow, from study/experimental design through to the data interpretation step. The training and teaching are carried out through various activities and approaches. The training aims to develop both cognitive and emotional domains of Bloom's taxonomy levels. Assessments and progress evaluations are conducted through different forms, some of which include (i) regular progress meetings (one-on-one and in a group), (ii) group discussions, (iii) progress reports (twice a year), and (iv) a

15-min oral presentation to the Department (at least twice a year), during which the student presents the progress on his/her work, concepts learned and applied, and receives constructive feedback through a Q&A session. This training/teaching is carried out through different activities and tasks: studying the literature, continuous one-to-one and group discussions, and demonstrations.

The Metabolomics South Africa (MSA), an affiliate to the Metabolomics Society, organizes a series of workshops annually with both introductory and advanced workshops that cover different aspects in the metabolomics workflow. These workshops are presented by senior SA scientists involved in metabolomics research or international collaborators (mostly leading scientists in the field). Students are encouraged to attend and actively participate in these workshops. Some of these workshops provide hands-on training; for example, in data processing, data mining, spectral pre-processing and feature annotation/identification. The PI also invites overseas collaborators (experts in the field) to be involved in research projects (Master's and/or PhD's), and interact with the students, training them on particular aspects, for example, the use of computational tools for metabolite annotation via online platforms. Outside of the core teaching and training students are encouraged to attend international webinars (for example, those delivered by the Metabolomics Society Early Members Network). To assist students who are struggling in grasping certain concepts or practical aspects the PI assigns a 'mentor', these mentors are often senior students who have already demonstrated advanced knowledge and skills. Throughout the year, the PI pays attention to (each) individual student, and sets regular progress meetings, providing guidance and additional training/teaching where necessary. To improve presentation and communication skills students are provided with opportunities to attend and present at conferences (local and international, and both poster or oral presentation); and encouraged to develop and form networks with their peers or senior scientists in the field. At the end of their master's and PhD programme, the student should have at least one research article published; and the student submits a dissertation (master's) or a thesis (PhD), which is evaluated by external assessors (generally, experts in the field).

5.3 Delivering computational education in a taught postgraduate university course via face-to-face teaching and distance learning

Edith Cowan University (Author SR) offers a Master of Bioinformatics programme. This programme includes a course called Mass Spectrometry in Systems Biology that covers metabolomics and proteomics spectral pre-processing, feature annotation/identification, and quality assessment/data

cleaning processes. This course has been offered yearly since 2020.

The learning outcomes are to (1) apply the theory of high-resolution mass spectrometry to metabolomics and proteomics by (2) critically analysing methods and guidelines on the above-mentioned tasks, (3) performing the above-mentioned tasks, and (4) defending decisions that informed performing the tasks. The learning outcome verbs align to Bloom's levels 3–5 (apply, analyse, evaluate). The course was designed to accommodate students from diverse discipline backgrounds (including no prior chemistry knowledge) and is offered via a dual delivery pattern (both on-campus and online enrolment modes are offered). Students are required to use spectral processing software to achieve the learning outcomes. This, in turn, required computers with sufficient compute power, software to be installed, transfer of large raw spectral files, and the ability to accommodate off-campus software use. Microsoft Azure Labs was used to overcome computational challenges. A virtual machine (VM) template was created, all software was pre-installed, and all files were uploaded. The template was cloned to create an identical private VM for each student. Students were able to access their VM from anywhere and at any time, and they could save their work each week. Thermo Fisher Scientific's Compound Discoverer was used for the spectral processing software. It reduced cognitive load by providing (1) a graphical-user interface (GUI), (2) visually intuitive presentation of concepts such as LC–MS peaks and MS/MS spectral matching, (3) an all-in-one software package to achieve all required content topics, and (4) the ability to process workflow steps in isolation.

To accommodate the two parallel study modes, the flipped classroom approach was used. Students watched a series of short (15-min) lecture recordings and completed required readings ahead of the 3-h weekly on-campus workshop. On-campus students participated in student-centred learning activities (such as Put Yourself on the Line mentioned earlier). Recordings were made of software demonstrations and relevant activities, then posted on the learning management system (LMS) for online students. Weekly discussion board activities were used to build inclusion between on-campus and online students, enable students to extend their own learning, and to promote critical thinking of metabolomics topics. Engagement in the discussion board activities was assessed as part of the portfolio assessment (see below) and offers an opportunity for formative feedback throughout the course.

Learning was assessed through two assessments. (1) Upon completing the content, students had an opportunity to explore and analyse how spectral processing, metabolite ID, and quality assessment/data cleaning are reported in the literature. In a 10-min oral presentation, students presented a published LC–MS metabolomics study, and critically

assessed the reported workflow. This assessment piece was mapped against learning outcomes #1 and 2, demonstrating up to the ‘analysis’ level of Bloom’s taxonomy. (2) In a summative portfolio, students completed, presented, and reflected on their own spectral processing, metabolite ID, and quality assessment/data cleaning workflow. The portfolio mapped against learning outcomes #1, 3, and 4, demonstrating Bloom’s levels 2–5 (understand through evaluate).

In summary, contemporary research-focused educational criteria can pose unique design and delivery challenges. Research and business solutions can help to overcome these challenges and support learning.

5.4 The application of asynchronous online course delivery in the training environment

The Birmingham Metabolomics Training Centre (Authors CW, WD) developed the small private online course (SPOC) to provide access to online courses for 50–100 trainees per course run. From 2017 to June 2021 a total of 384 trainees participated in the courses. These courses were typically operated twice per year though were operated more frequently in 2020 and 2021 as the desire for online training increased during the Covid-19 epidemic.

Learning objectives were designed across all levels of Blooms cognitive domain to remember and understand knowledge and to apply and evaluate data processing/analysis methods that are applied in metabolomics. An asynchronous online format was used to deliver the course. This allowed a high number of trainees to attend the courses; typically, 30–100 trainees were present on each course. The course was delivered using a professional online course platform (FutureLearn) with multiple delivery tools. Each week focused on specific learning objectives and included a variety of delivery methods (short videos provided by the trainer, articles, exercises, polls and quizzes to allow the trainees to assess their knowledge). The training format applied both teacher-centred and student-centred learning. Teacher-centred activities included short video lectures whereas, student-centred tasks included workshops with step-by-step protocols to perform the analysis of data. Courses were operated over 3 or 4 weeks with an estimated learning times of 4 h per week. Background reading was provided for those who did not have the background knowledge of key concepts, however, to undertake all background reading would require a greater time commitment than 4 h per week. Trainees retained access to the course material after the course finished so could revisit the course material at any time. Links to papers published in the scientific literature were included to allow trainees to observe experimental processes and research performed by other experts in metabolomics.

The application of this asynchronous training method allows participants to complete the course when they have time available. It is advantageous to complete within the designated weeks to receive support from the trainers and fellow participants but is not essential. The low time commitment (4 h per week) also allows participants without the required background knowledge to undertake extra reading during the course. The use of technology enhanced the training experience, in particular discussion forums supported trainer-to-trainee and trainee-to-trainee interaction and sharing of knowledge. In addition, a one-hour live Q&A session was operated via the Zoom platform at a time most suitable for the trainees. Trainees were encouraged to post questions for the Q&A from Week 1 of the course. The most frequently asked questions and those deemed as important by the trainers were answered in an end-of-course video provided by one of the trainers. If all questions could not be answered during the live session, questions were answered on the discussion boards. The Q&A was also recorded and uploaded to the FutureLearn platform for those who could not join live or to re-visit the discussion. Self-assessment steps were included via quizzes or tests, if an incorrect answer was submitted participants were direct back to particular sections of the course to review the course material. Participants’ completion of the course tasks was logged in the online platform and a Certificate of Achievement was provided for those who completed the course.

In summary, the SPOC courses provided flexibility in the provision of training materials, flexibility in the time when trainees access the training materials, online support of trainees and the ability to train many trainees.

6 Summary

The field of metabolomics is complex, multidisciplinary, and not standardized. Accordingly, researchers working in metabolomics must possess in-depth cross-disciplinary *content knowledge*, a high level of subjective experience-based *judgement knowledge*, and the associated complementary skill sets. To develop this knowledge-base and skill sets necessitates appropriate training and educational initiatives within the metabolomics community. Initiatives must aim to be effective, but also feasible. As such, both pedagogical best practice and metabolomics-specific contextual challenges must be considered with equal weight. In this paper, we have presented consolidated pedagogical guidance for educators and trainers in metabolomics including considerations for addressing metabolomics-specific challenges. We provided case studies from our own practice to illustrate how this can be achieved. As metabolomics research grows in importance and diversity it is essential that best practice guidelines are put in place to ensure maximally effective

learning for future metabolomics scientists. Although no community currently exists to discuss best practices and develop training materials, the authors recommend that a community focused on metabolomics teaching and training is developed and the Metabolomics Society should lead in such an effort.

Acknowledgements The authors wish to thank Michelle Reid for her valuable feedback as an end-user. SNR wishes to thank Ralf Tautenhahn (Thermo Fisher Scientific) for licensing and technical support in using Compound Discoverer and Paul Haskell-Dowland (Edith Cowan University) for technical support in using Microsoft Azure Labs in her teaching practice. The authors wish to thank the Medical Research Council (MRC) in the UK for funding CW and WBD (MR/S010483/1).

Author contributions All authors contributed to conceptualising and writing the manuscript. All authors read and approved the final manuscript.

Data availability Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

Declarations

Conflict of interest The authors have no disclosures of potential conflicts of interest related to the presented work.

Research involving human and/or animal participants No research involving human participants or animals was performed for this work.

References

- Anderson, L. W., Krathwohl, D. R., & Bloom, B. S. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives* (Abridged). Longman.
- Beger, R. D., Dunn, W., Schmidt, M. A., Gross, S. S., Kirwan, J. A., Cascante, M., Brennan, L., Wishart, D. S., Oresic, M., Hanke-meier, T., Broadhurst, D. I., Lane, A. N., Suhre, K., Kastenmuller, G., Sumner, S. J., Thiele, I., Fiehn, O., Kaddurah-Daouk, R., for "Precision Medicine and Pharmacometabolomics Task Group"-Metabolomics Society Initiative. (2016). Metabolomics enables precision medicine: "A White Paper, Community Perspective." *Metabolomics*, 12, 149.
- Biggs, J. (1996). Enhancing teaching through constructive alignment. *Higher Education*, 32, 347–364.
- Bloom, B. S., et al. (1956). Taxonomy of educational objectives: The classification of educational goals. In B. S. Bloom (Ed.), *Handbook 1, Cognitive domain; by a committee of college and university examiners* (1st ed.). London: Longman Group.
- Boyce, M. C., Lawler, N. G., Tu, Y., & Reinke, S. N. (2019). Introducing undergraduate students to metabolomics using liquid chromatography-high resolution mass spectrometry analysis of horse blood. *Journal of Chemical Education*, 96, 745–750.
- Brennan, L. (2017). The nutritional metabolomics crossroads: How to ensure success for dietary biomarkers. *American Journal of Clinical Nutrition*, 105, 293–294.
- Bundy, J. G., Davey, M. P., & Viant, M. R. (2008). Environmental metabolomics: A critical review and future perspectives. *Metabolomics*, 5, 3.
- Fiehn, O., Robertson, D., Griffin, J., van der Werf, M., Nikolau, B., Morrison, N., Sumner, L. W., Goodacre, R., Hardy, N. W., Taylor, C., Fostel, J., Kristal, B., Kaddurah-Daouk, R., Mendes, P., van Ommen, B., Lindon, J. C., & Sansone, S.-A. (2007). The metabolomics standards initiative (MSI). *Metabolomics*, 3, 175–178.
- FindAMasters. (2022). <https://www.findamasters.com>. Accessed 1 January 2022
- Fraser, P. D., Aharoni, A., Hall, R. D., Huang, S., Giovannoni, J. J., Sonnwald, U., & Fernie, A. R. (2020). Metabolomics should be deployed in the identification and characterization of gene-edited crops. *The Plant Journal*, 102, 897–902.
- Haug, K., Salek, R. M., Conesa, P., Hastings, J., de Matos, P., Rijnbeek, M., Mahendraker, T., Williams, M., Neumann, S., Rocca-Serra, P., Maguire, E., González-Beltrán, A., Sansone, S.-A., Griffin, J. L., & Steinbeck, C. (2012). MetaboLights—An open-access general-purpose repository for metabolomics studies and associated meta-data. *Nucleic Acids Research*, 41, D781–D786.
- Hollywood, K. A., Schmidt, K., Takano, E., & Breiting, R. (2018). Metabolomics tools for the synthetic biology of natural products. *Current Opinion in Biotechnology*, 54, 114–120.
- Hoque, M. (2017). Three domains of learning: Cognitive, affective and psychomotor. *The Journal of EFL Education and Research*, 2, 45–51.
- Jarvis, P. (2012). *Adult education and lifelong learning: Theory and practice*. Taylor and Francis.
- Kell, D. B., & Oliver, S. G. (2016). The metabolome 18 years on: A concept comes of age. *Metabolomics*, 12, 148.
- Krathwohl, D. R., Bloom, B. S., & Bertram, B. M. (1964). *Taxonomy of educational objectives: The Classification of educational goals. Handbook II: Affective domain*. Longmans.
- La Frano, M. R., Amin, S., & Fanter, R. K. (2020). Participation of undergraduate students in a controlled feeding study with metabolomics analysis to enhance learning of metabolism. *Journal of Chemical Education*, 97, 1595–1603.
- Mashego, M. R., Rumbold, K., De Mey, M., Vandamme, E., Soetaert, W., & Heijnen, J. J. (2007). Microbial metabolomics: Past, present and future methodologies. *Biotechnology Letters*, 29, 1–16.
- Mendez, K. M., Pritchard, L., Reinke, S. N., & Broadhurst, D. I. (2019). Toward collaborative open data science in metabolomics using Jupyter Notebooks and cloud computing. *Metabolomics*, 15, 125.
- Mulder, N., Schwartz, R., Brazas, M. D., Brooksbank, C., Gaeta, B., Morgan, S. L., Pauley, M. A., Rosenwald, A., Rustici, G., Sierk, M., Warnow, T., & Welch, L. (2018). The development and application of bioinformatics core competencies to improve bioinformatics training and education. *PLoS Computational Biology*, 14, e1005772.
- Oliver, S. G., Winson, M. K., Kell, D. B., & Baganz, F. (1998). Systematic functional analysis of the yeast genome. *Trends in Biotechnology*, 16, 373–378.
- Print, M. (1993). *Curriculum development and design* (2nd ed.). Allen & Unwin.
- Rai, A., Yamazaki, M., & Saito, K. (2019). A new era in plant functional genomics. *Current Opinion in Systems Biology*, 15, 58–67.
- Ratray, N. J. W., Deziel, N. C., Wallach, J. D., Khan, S. A., Vasiliou, V., Ioannidis, J. P. A., & Johnson, C. H. (2018). Beyond genomics: Understanding exposotypes through metabolomics. *Human Genomics*, 12, 4.
- Sandusky, P. O. (2017). Introducing undergraduate students to metabolomics using a NMR-based analysis of coffee beans. *Journal of Chemical Education*, 94, 1324–1328.
- Sud, M., Fahy, E., Cotter, D., Azam, K., Vadivelu, I., Burant, C., Edison, A., Fiehn, O., Higashi, R., Nair, K. S., Sumner, S., & Subramaniam, S. (2015). Metabolomics Workbench: An international repository for metabolomics data and metadata, metabolite standards, protocols, tutorials and training, and analysis tools. *Nucleic Acids Research*, 44, D463–D470.
- Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, 12, 257–285.

- Tinte, M. M., Chele, K. H., van der Hooft, J. J. J., & Tugizimana, F. (2021). Metabolomics-guided elucidation of plant abiotic stress responses in the 4IR Era: An overview. *Metabolites*, *11*, 445.
- Wang, M., Carver, J. J., Phelan, V. V., Sanchez, L. M., Garg, N., Peng, Y., Nguyen, D. D., Watrous, J., Kapon, C. A., Luzzatto-Knaan, T., Porto, C., Bouslimani, A., Melnik, A. V., Meehan, M. J., Liu, W. T., Crusemann, M., Boudreau, P. D., Esquenazi, E., Sandoval-Calderon, M., ... Bandeira, N. (2016). Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking. *Nature Biotechnology*, *34*, 828–837.
- Weber, R. J., Winder, C. L., Larcombe, L. D., Dunn, W. B., & Viant, M. R. (2015). Training needs in metabolomics. *Metabolomics*, *11*, 784–786.
- Zanetti, K. A., Hall, R. D., Griffin, J. L., Putri, S., Salek, R. M., Styczynski, M. P., Tugizimana, F., & van der Hooft, J. J. J. (2019). The metabolomics society-current state of the membership and future directions. *Metabolites*, *9*, 89.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.