SECTION 3

CHARACTER TAUGHT:

PERSPECTIVES ON EVALUATING CURRICULUM STRATEGIES AND ACTIVITIES

This section starts with a more general overview on the nature and use of self-evaluation methods and instruments. It looks at the different approaches and methods that could be used to self-evaluate formal or non-formal curriculum interventions/strategies/activities designed to develop character. There are sub-sections on the use of surveys and self-report scales, observation of lessons, group interviews and dilemmas.

Each sub-section provides perspectives on different methods that can be used by schools to self-evaluate their character education curriculum strategies and activities. The sub-sections have been written by individuals with experience in using the various methods. It is important that schools adapt these methods and suggested instruments to best meet their own needs and contexts.

3.1 PLANNING AN EVALUATION

3.1

A broad overview of what needs to be thought about when developing an evaluation of a character education activity or intervention is provided, introducing the importance of understanding what you are trying to evaluate, along with how the context of the evaluation may affect its design.

The Centre for Youth Impact

3.1.1 Where to start: what questions would you like to answer?

All effective evaluations start with a research question. It is important to take the time to reflect and consider the question/s you'd like to answer before beginning the data gathering. The risk in doing this is the discovery at a later stage that the information needed is not available, and relevant conclusions cannot be drawn.

In planning a self-evaluation, it is important to think about the kind of information to be collected, and the intended use: the questions you want to answer.

Generally there are two kinds of evaluation – **process** and **impact**.

A **process** evaluation explores the way in which a character education intervention[1] has been implemented, to provide useful information on the day-to-day experience of setting up and running the project. Typical questions are:

- What did the intervention intend to achieve?
- Did the intervention go ahead as intended?
- Were some things harder than expected?
- Were some things easier?
- Was the intervention delivered consistently?
- Was it well received by students?

An **impact** evaluation assesses the change that has been brought about as a result of the project. Typically, this kind of evaluation tries to:

- identify the state of play at the start (the baseline);
- compare this to the state of things at the end (the endline); and
- compare this change over time to a situation in which the intervention did not occur (the counterfactual).

An impact evaluation aims to pinpoint whether a change has occurred under a specific set of circumstances. This is often called taking an experimental, or quasi-experimental, approach. This means that the situation is controlled, as in a laboratory, in order to try and attribute a particular change – improved attentiveness for instance – to one intervention. The 'gold standard' of this kind of research is the Randomised Controlled Trial (RCT). As has already been noted in section 1, this is a difficult kind of study to conduct and requires considerable expertise and resource. It is likely to be beyond the means of most teachers and most schools to carry out an RCT to assess the effectiveness of a character education project.

[1] We use the term 'intervention' to denote the activities that are the focus of the evaluation. Others may refer to 'projects', 'programmes', 'services' or simply 'activities'.

It is likely that as someone working on character education you will, in any event, want to know the answers to both types of questions: process and impact. You will want to know something about how a project went, and whether or not it was effective. A project may seem to go well from start to finish, but if it had no effect, this seems less significant. At the opposite extreme, a project may have been very challenging to implement, but if it can be demonstrated that it has had a real impact, then the pain was worthwhile.

It is likely therefore that the most useful evaluation for this context will be a blend of approaches that provides useful evidence to answer both process and impact questions. Useful evidence will be:

- Reliable;
- Relevant; and
- Reasonable.

This means that evidence collected will be as reliable as can be expected, as relevant as it possibly can be to the school, its staff and its students, and reasonable to gather and analyse without making undue demands on staff and student time and resource.

3.1.2 The context

What questions can be answered in the school setting?

It is important to understand the limitations of evaluating the development of character in a school setting. Although children and young people spend a considerable amount of time at school, they are exposed to a wide variety of influences outside the school setting which will have a significant impact on their character development.

By way of comparison, we might think of children learning language.

Usually, picking up language involves social exposure through childhood. Formal schooling is not necessary. Television and radio cannot act as a substitute for human interaction, but such interaction can be indirect. Children can pick up a language from hearing it spoken amongst adults who are not addressing them directly (Smith, 2016). Children can even develop a language in the absence of normal interaction with adults if they are in a group setting with other children (Reed Magazine, 2013).

Assuming regular exposure to language from more than one adult, even in the absence of formal schooling, children will develop a full command of their mother tongue. They may indeed develop more than one language in this way.

This does not mean that formal schooling, learning to read and write and acquiring further knowledge will not have effects on vocabulary, grammar and expression. It does mean, however, that children who are socialised in ordinary ways have a remarkable capacity to develop language without formal teaching. Any input from a school is always building on this innate capacity. Schooling and education play a major role in helping to build a child's innate capacity to develop language.

This has implications for the assessment of language development in the classroom context. There is a long history of literacy assessment, and not surprisingly much of that assessment is based on gauging development without attempting to pin down attribution of that development to the school itself, except by proxy through league tables and so on. It is possible to attribute language gains to specific programmes only in very circumscribed situations; on the whole it is a difficult thing to do (Ofsted, 2013).

Character development is similar to language development. A student's character will be influenced by all of their social interactions in and outside the school. Attributing change to an in-school methodology only is therefore complex and, even if relevant, is likely to fail the reliable and reasonable tests.

Instead, it is a good idea to focus research questions more narrowly and to understand an evaluation of character education as a form of reflective practice that seeks to ask questions, collect information and build meaning in order to continually improve practice.

Figure 2: A step by step guide to the self-evaluation of a character education intervention

STEP 1: DESIGN THE INTERVENTION

STEP 2: IDENTIFY EVALUATION QUESTIONS

STEP 3: CHOOSE AND APPLY A METHOD

STEP 4: ACT ON THE FINDINGS

STEP 1: DESIGN THE INTERVENTION

This is about looking at the available evidence to inform the design of the taught character education intervention. The design of the intervention should respond to an identified need.

Things to consider when designing the character education intervention:

- what evidence is there to inform the design of the intervention?

  Key sources of evidence include:
  The Early Intervention Foundation http://www.eif.org.uk/
  The Education Endowment Foundation https://educationendowmentfoundation.org.uk/
  The Education Datalab http://educationdatalab.org.uk/
  Investing in Children, a catalogue of evidence-based programmes http://investinginchildren.eu/

  The Centre also has a number of different resources available, including:
  Knightly Virtues www.jubileecentre.ac.uk/knightlyvirtues
  My Character www.jubileecentre.ac.uk/mycharacter
- is there the opportunity to involve students in the design of an intervention? Would this support engagement with the intervention?
- is the intervention intended to be directed to particular groups of students?
- how can you build in evaluation into the design of the intervention?

STEP 2: IDENTIFY THE EVALUATION QUESTIONS

Thinking about the questions that will direct the evaluation will help to determine which research methods will be used. If there is an interest in the impact of the intervention, or how the intervention has been introduced, and whether any changes have been made to the original model (also known as implementation fidelity), this will influence the questions to be asked.

Things to consider when identifying the evaluation questions:

- is there an interest in the impact and outcomes of a character education intervention (impact evaluation) or in how the intervention has been put in place (process evaluation)?
- should the evaluation also audit how much time staff spent on both developing and delivering the intervention?
- if the evaluation is focused on implementation fidelity, is it relevant to consider the following:
  - changes to the group targeted by the intervention?
  - changes in the delivery of the intervention? For instance, changes in the content of the intervention, changes in the delivery of the intervention, changes in the intensity of the intervention, or changes in the timing of the delivery.

STEP 3: IDENTIFY THE RESEARCH METHODS TO BE USED TO COLLECT AND ANALYSE THE DATA

Choosing appropriate research methods will help you to collect data appropriate to the evaluation question. Section 3 provides more information on the methods that could be used.

Things to consider when identifying research methods:

- will the research methods capture the relevant information?
- how much time and resource are available to dedicate to the collection and analysis of data?

STEP 4: ACT ON THE FINDINGS

An evaluation will have the broad aim of improving character education provision within your school. Thinking about how to feedback the findings of the research could help you to identify how the findings could have the most impact and support the development of character education provision in future.

Things to consider when acting on the findings:

- what opportunities will there be to act on your findings?
- how will colleagues be informed about the findings?
- will those who participated in the evaluation be informed of the findings? How will this be done?

3.1

CHARACTER TAUGHT

3.1.3 Summary: advice about evaluating character education interventions to discover 'what works'

- When planning a character education intervention, thinking about how to build in an evaluation from the very start can help to schedule the evaluation into a busy school year.
- Related to the above point, think about the questions the character education evaluation should answer. It may be that there is an interest in the impact of an intervention or how the intervention has been implemented and what effect this may have had on possible impact.
- An evaluation plan will never be perfect and will always be limited by its context. It is important to be aware of the limitations of the evaluation and that students' character will always be affected by wider influences. There is a need to be aware of these wider factors so that an intervention is not incorrectly attributed with causing an effect. Reviewing the evaluation will help to identify some of these limitations and will also help inform future evaluations.
- Think about how feedback of the results of your evaluation will be given to those people who contributed to it.

References

Ofsted (2013) *Improving Literacy in Secondary Schools: A Shared Responsibility*. Available at https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/413182/Improving_literacy_in_secondary_schools.pdf (Accessed: 18 August 2016)

Reed Magazine (2013) *Birth of a Language*. Available at http://www.reed.edu/reed_magazine/june2013/articles/features/signlanguage/signlanguage1.html (Accessed: 18 August 2016)

Smith, N. (2016) *Chomsky, Ideas and Ideals (third edition)*. Cambridge: Cambridge University Press

3.2 SURVEYS

The design and administration of surveys are considered here, along with looking at approaches to the analysis of the data generated by a survey. The limitations of the approach are also discussed, along with the need to use this approach in combination with other methods. The contribution that logic models can make to inform the structure of an evaluation is also considered.

Ralph Scott, Demos

3.2.1 Introduction

Of all the instruments in the social researcher's toolbox, surveys are probably the most vividly recalled in the popular imagination – the source of many column inches in newspapers each day. Used appropriately, they enable the approximate quantification of an often complex social reality – translating life into numbers, and therefore simplifying to aid interpretation. But for this very reason they should be used and interpreted with caution – particularly in evaluating changes in character, a discipline which is still in its infancy and may always be difficult to assess through self-report instruments, in general, and questionnaires, in particular, as discussed elsewhere in this handbook.

This section explains the considerations to be borne in mind when using surveys as part of an evaluation of a character education intervention. Before doing so, it is worth reiterating the broader point on triangulation made earlier on: while surveys are often used to track self-reported character strengths and weaknesses, one should also consider monitoring change in behavioural measures (such as attendance, behaviour and academic progress) and remember that wider developmental progress might not be evidenced by self-report questionnaires. Instead, this can often be picked up through qualitative work: session observations and interviews with the participants, their teachers and others with experience of the programme.

Therefore the character outcomes identified through the survey should not form the sole basis of your evaluation, as they may not capture everything you are interested in.

This sub-section draws on recent evaluations carried out by Demos of the Department for Education's Character Grant, particularly the evaluation of The Scout Association's Character by Doing programme (Demos, 2016). Readers interested in discovering more about self-evaluation are advised to read the Education Endowment Foundation's (2013) accessible DIY evaluation guide (focused on character to enhance attainment, but includes useful advice on method), and those looking for a good grounding in social research methods could do worse than pick up a copy of Bryman's (2012) textbook of that same name. A further example of a survey which may be of interest is the Values in Action (VIA) youth survey designed for those aged 10-17 years old. This is explained further in sub-section 4.2 Ideas and activities to encourage student self-reflection on character.

3.2.2 Research objectives and design

It may seem obvious, but before starting to design a survey or engaging in any aspect of the research, it is important to clearly set out the research objectives, to direct all of the research activities at answering a particular question or set of questions. This should prevent sloppy research that fails to address a crucial question you are interested in, or wasting time on activities which are not useful.

This could be as simple as spelling out what it is intended to find out through the research – for example, for the Character by Doing evaluation, the research question was as follows:

What impact does participation in non-formal learning, in the form of a six-month programme modelled on participation in Scouting, have on character and academic outcomes? (Demos, 2016)

A slightly more sophisticated approach is to develop a 'logic model' for the programme – laying out on one page the context, inputs, outputs, and intended outcomes and impacts of the intervention. This helps the evaluator to better understand what the intervention is hoping to achieve, and therefore what measures to include in the

Figure 3: Example of Character by Doing logic model

CONTEXT	INPUTS	OUTPUTS	
		Activities	Participation
What is the issue? Young people require a range of character capabilities to live happy and successful lives The education system is currently struggling to develop these in all young people Non-formal learning is known to have positive impacts on character, educational and wellbeing outcomes Too few young people are participating in non-formal learning like Scouting There is also unequal access to non-formal learning through schools and a perceived lack of diversity	**What do we need to achieve our goals?** *Place:* 6 schools (4 in South East; 2 in East Midlands) *Delivery personnel:* - 6 staff from The Scout Association or Regional Services Team - approx. 20 teachers or teaching assistants - approx. 21 volunteers *Time:* 6 months (Sep 2015 – Mar 2016) *Partners:* schools, local Scout groups, DfE, community organisations *Venues:* schools, local community, Scout activity centres *Resources:* £5,000 incentive for schools, start-up boxes, tents etc	**What do we have to do to ensure our goals are met?** Young people will participate in a programme modelled on the Cub programme *Activities:* - approx 24 weekly two-hour sessions during term-time - overnight experience - community action project (Our World challenge badge) *Deliverers will:* - receive training in early September - attend two forum events during course of project	**Who needs:** ✳ **to participate?** ✳ **to be involved?** ✳ **to be reached?** *Number:* approx. 120 participants across 6 schools *Age:* 8–10-year-olds *Target groups:* Pupils receiving Free School Meals (FSM), Black, Asian, and minority ethnic (BAME) pupils, not current Scouts *Recruitment:* voluntary and compulsory depending on school *Wider engagement:* parents, local Scout community, SLT and school community, neighbouring schools

research design. To illustrate what is meant by a logic model, an example using the Character by Doing (Demos, 2016) programme is provided below. Ideally, this would be developed in consultation with those involved in the intervention – particularly those helping to design and deliver it, but also potentially those participating.

Once a research question is set, it is time to decide on methods to help address the question. The purpose of this sub-section is to discuss the use of surveys as a method, and so the remainder will focus on this question – however, the point on triangulation should be remembered at all times by those leading the evaluation.

The most popular use of surveys in evaluating programmes tends to be what is called a 'pre- and post-' design: meaning that those participating are asked to complete a survey prior to taking part, and then to take the same survey on completion of the programme. The key outcome of interest is the change in intervention participants' average reported scores following participation – essentially showing whether during the course of the intervention there has been a positive or negative change in empathy, for example. Without using a pre- and post-design it is difficult to show change over time, and therefore determine the effectiveness of any intervention.

OUTCOMES: Learning	OUTCOMES: Action	IMPACTS
What do we think the participants will know, feel, or be able to do after participation in the programme?	**How do we think the participants will behave or act differently after participation in the programme?**	**What kind of impact can result if participants behave or act differently after participation in the programme?**
In line with the Youth Outcomes Framework, participants may see improvements related to: communication, confidence and agency, managing feelings, problem solving, relationships and leadership, grit and resilience Improved attitudes towards self, others, education and community Participants may see improvements on educational attainment and progress	Participants demonstrate more commitment at school: attendance, autonomy, punctuality and behaviour Participants are more likely to engage in volunteering and service in their communities Participants remain in Scouting	Young people receive more rounded and character-building education: leading to better outcomes Non-formal education is validated in eyes of school leaders and policy makers. Improves teacher practice Programme is sustained and expanded locally: more young people can participate in Scouting Develops a model which other schools can use: further expands opportunity

3.2

3.2.3 Questionnaire design

Concepts

Before addressing in detail how to put together a questionnaire, some thought should be given to the intended outcomes to be achieved and subsequently evaluated through the programme. These are inextricably linked to the conceptual understanding of character. A widely used conceptual definition, set out at the beginning of this volume, is that provided by the Centre, which emphasises four main domains of character - moral, intellectual, civic and performance – placing a particular emphasis on the first of these. Other conceptual approaches more influenced by personality psychology are also prevalent, and tend to place a particular focus on performance characteristics, such as grit and resilience.

The conceptual approach to character which is adopted will influence the expected outcomes, and therefore the questions included in the questionnaire to evaluate these outcomes. The process of thinking about the aims of the programme (whether through designing a logic model or otherwise) should make clear the intended outcomes, which could be anything from propensity to volunteer to developing virtue knowledge.

Designing the questionnaire

Once the conceptual understanding of character, and the outcomes you wish to assess, are decided upon, the questionnaire can be developed.

This should be comprised of questions which accurately assess the outcomes of interest – often more than one question will be used for each outcome. When deciding what questions to use, by far the most preferable option is to use questions that have already been used by other researchers, or feature in government surveys. This has a number of advantages:

- **Confidence:** if a questionnaire has been used for research many times before, there is a degree of certainty that it is able to do what it says it does, and that any issues faced in administering it have been faced by someone else before – therefore advice will be out there.
- **Convenience:** it is also a lot easier to use questions from a pre-existing questionnaire than draft one entirely from scratch. Teachers are time-poor, and should use short-cuts where they can, especially if it improves quality!
- **Validity:** questions that have been used and carefully evaluated as part of government or academic surveys are more likely to be valid than completely new and untested questions - the question will more accurately assess the concept to be considered. Often researchers will have gone through a trial and error process, using statistical methods and cross-referencing to derive the best questions that measure what is intended to be measured.
- **Reliability:** such questions will often also have been found to be reliable, which means that responses from the same individual are consistent over time.
- **Comparability:** at sufficient scale, using questions that have already been used to evaluate other interventions gives a sense of the relative impact of the intervention, enabling placement of the observed change in context.

What is to be measured may sometimes be specific to the intervention, in which case original questions will need to be drafted. Even in this case, it is best practice to see what has gone before, and how existing questions can be adapted to suit the purpose. If substantial parts of someone else's survey are to be used it is good practice to seek their permission first.

There are a number of other things to bear in mind when drafting a questionnaire:

- think about the respondents: Who are they? What is their reading age? What is their attention span like? Are the questions age or culturally appropriate?
- a questionnaire should be as short as possible, while also getting all the information required – at primary age, anything that takes much longer than 5 minutes to complete (approx. three pages of A4) will need teacher support; for secondary, 10-15 minutes is recommended. The survey could be extended if it is built into a learning activity.
- ensure that identifying information is requested, such as full names and class name – this is crucial to tracking individual change before and after the programme. This complication could be avoided by looking at the group level, rather than individual level, change. If confidentiality is a concern, produce a code for each respondent and ask him/her or their teacher to include that on the questionnaire.
- if possible, pilot the questionnaire by getting another class to complete it – see what issues come up and adapt it accordingly.

Regardless of whether a pre-existing or original questionnaire is used, various types of question will arise. The most commonly used for evaluation of character is to ask respondents to reflect on their own strengths and weaknesses, and place themselves on a scale, as this enables researchers to quantify the change and thereby this could more accurately reflect that an aspect of character has changed.

For the DfE evaluations, use was made of the Behavioural Insights Team's (BIT, 2006) questionnaire, originally designed to evaluate the impact on character of various youth social action programmes. In order to estimate character outcomes, this asks respondents to score themselves on a 0-10 Likert scale across a range of different statements. The scores on these statements are then averaged for each respondent, giving an overall score for each 'construct' – that is the character outcome of interest. For example, for co-operation, the following questions are asked:

Please tell us how true the following statements are about you.
There is a scale where 0 = "not at all true" and 10 = "completely true".
Tick only one box on each line.

NOT AT ALL TRUE — **COMPLETELY TRUE**

	0	1	2	3	4	5	6	7	8	9	10
I enjoy working together with other students my age											
I can work with someone who has different opinions to me											
I am able to compromise and resolve differences of opinion											

The average score across these responses makes up the overall co-operation 'score'.

As this was aimed at teenage respondents, for the Character by Doing evaluation questions were adapted to be more suitable for our age group (Key Stage 2). This included reducing the overall number of questions – so that there was only one statement per character strength – and also radically simplifying the language and content of the questions.

Also, on the BIT questionnaire, some of the statements are 'reverse-coded', meaning that a high score corresponds negatively to the outcome: for example, a high score for the statement 'most of the time I don't want to go to school' contributes negatively to the 'school attitudes' outcome. This was found to confuse primary respondents, who could put low scores to indicate negative feeling, rather than a negative response to the statement.

Some examples of the primary statements – drawn from pre-existing validated questionnaires – are below:

- **Empathy:** I feel bad when somebody gets their feelings hurt
- **Leadership:** I feel comfortable being a group leader
- **Communication:** I am good at explaining my ideas to other people

While using longer Likert scales is good for evaluation – as it allows for finer-grained measurement of change, e.g. from a 7 to an 8 – for primary age it is recommended to use simpler scales, and to bring these to life through use of pictures and colour.

Another important consideration relates to who is answering the questionnaire – while it might seem appropriate to focus only on the respondents themselves, particularly for younger respondents, it is sometimes useful to triangulate those findings with teacher assessments. An example of this approach would be Goodman's Strengths and Difficulties questionnaire (freely available online) (Youth in Mind, 2016), which in addition to a self-report questionnaire, includes a version for both teachers and parents to complete on behalf of a young person. Some examples of each are below and can be found via the following website www.sdqinfo.org.

Figure 4: Strengths and difficulties questionnaire – student questionnaire

	NOT TRUE	SOMEWHAT TRUE	CERTAINLY TRUE
I try to be nice to other people. I care about their feelings.	☐	☐	☐
I am restless. I cannot stay still for long.	☐	☐	☐
I get a lot of headaches, stomach-aches or sickness.	☐	☐	☐
I usually share with others (food, games, pens etc.).	☐	☐	☐
I get very angry and often lose my temper.	☐	☐	☐
I am usually on my own. I generally play alone or keep to myself.	☐	☐	☐
I usually do as I am told.	☐	☐	☐
I worry a lot.	☐	☐	☐

Figure 5: Strengths and difficulties questionnaire – teacher questionnaire

	NOT TRUE	SOMEWHAT TRUE	CERTAINLY TRUE
Considerate of other people's feelings.	☐	☐	☐
Restless, overactive, cannot stay still for long.	☐	☐	☐
Often complains of headaches, stomach-aches or sickness.	☐	☐	☐
Shares readily with other children (treats, toys, pencils etc.).	☐	☐	☐
Often has temper tantrums or hot tempers.	☐	☐	☐
Rather solitary, tends to play alone.	☐	☐	☐
Generally obedient, usually does what adults request.	☐	☐	☐
Many worries, often seems worried.	☐	☐	☐

One risk to bear in mind with this is the additional teacher workload it entails – therefore it is a decision not to be taken lightly!

It is important to note that in all of these examples, within the context of the questionnaire, it is unclear to the respondent how the statements are subsequently coded – these statements are mixed in with others assessing other outcomes. The idea is that this gives a 'truer' account of their character, aiming to avoid desirability bias by not making it clear which responses are the 'good' ones – although it makes it more difficult for the respondent to reflect on their own character.

The original BIT questionnaire includes a style of question that goes further in its attempt to overcome desirability bias – by testing the willingness of the respondent to actually volunteer, having taken part in a youth social action project. The question appears as a simple tickbox, asking whether the respondent would like to hear more about future volunteering opportunities – however, this is in fact an assessment of whether their stated willingness to continue volunteering elsewhere in the questionnaire transpires in reality. This kind of question comes with ethical concerns – and should in all likelihood only be used where the opportunities being advertised are in fact available.

3.2.4 Formative assessment

It is also possible to pursue an approach more resembling 'formative assessment' (for a definition see sub-section 1.3.2), whereby how the questions correspond to outcomes is more transparent to those completing the questionnaire, and respondents reflect on these more regularly, rather than through a pre- and post-design. This encouragement of self-reflection is thought to help with character development in its own right, and is perhaps a more appropriate method of evaluating a whole-school or ethos driven approach to character development. It does invite the possibility of bias however, whereby respondents give the responses they think are expected of them, as opposed to what they might do 'unobserved'.

Perhaps the best known example of this is the approach taken by the Knowledge is Power Programme (KIPP) schools in the USA. They encourage students and teachers to report on an individual's character strengths using the Character Growth Card, developed by Character Lab, at various points throughout the academic year. It can be used as a means of assessing progress and differences in self and teacher-assessment, start a conversation about character and to set goals together. As they put it:

> **It's important to note that this tool should not be used to diagnose or compare children, nor to compare schools or programs. Please use it to help children focus on their own growth and development in these areas, and as a positive conversation starter.**
>
> **Character Lab, 2015**

Examples of the kinds of statements (rated on a scale of 1 to 7), and the character strengths they correspond to are provided in the table below:

Figure 6: Examples of character strengths and related statements

Character strength	Statements
Gratitude	Did something nice for someone else as a way of saying thank you.
	Expressed appreciation by saying thank you.
	Showed appreciation for opportunities.
	Recognised what other people did for them.
Optimism	Believed that effort would improve his/her future.
	When bad things happened, s/he thought about things they could do to make it better next time.
	Stayed motivated, even when things didn't go well.
	Believed that s/he could improve on things they weren't good at.
Curiosity	Was eager to explore new things.
	Asked questions to help him/her learn better.
	Took an active interest in learning.

3.2.5 Questions about the implementation of the intervention

Beyond questions based on statements and scales, provision of open response questions should be considered, which give respondents the freedom to express themselves in their own words. This can help to avoid the respondent being led by the statements and scales used in closed questions. Although answers to these questions are inevitably more difficult to analyse, more may be learned from the rich, qualitative responses than from the self-reported scores. In terms of moral reasoning, there is also a great deal to be learned from moral dilemma questions, which are covered in the next section of this handbook.

In the post-intervention survey, questions about the intervention itself could be asked. These types of questions are considered less useful in assessing impact, as they are at risk of what is known as desirability bias – whereby respondents respond positively to questions such as ‘taking part in this programme has made me confident’ not because they are significantly more confident, but because they enjoyed the programme and want to reflect that in their answer.

Questions about the intervention can be very useful however in assessing the success of the programme in itself, and identifying strengths and weaknesses and room for improvement. For this reason, open response questions can be useful, as it can sometimes be difficult to predict how respondents will reflect on their experiences. Some examples of post-programme statements from our evaluation of the Character by Doing programme are as follows:

- I have enjoyed doing Cubs;
- I would recommend Cubs to a friend;
- I look forward to Cubs every week;
- I want to keep doing Cubs.

3.2

3.2.6 Sample and administration

A very important consideration when carrying out an evaluation is sampling – deciding who completes the questionnaire. In the case of a school-based intervention it seems obvious that every respondent should complete the questionnaire – however, who ends up taking part in the programme could influence the results, particularly if participation is voluntary. For example, a sport intervention might appeal more to boys – therefore the group might end up with more boys than girls, which due to factors external to the intervention could make it appear as though there has not been much change in character.

In many ways, the pre- and post-design – which takes account of individuals' starting points when reporting progress – helps to account for these differences. But it is always worth tracking and reporting on things like the gender, age and ethnic breakdown of the intervention group, as well as monitoring data such as free-school meal eligibility, special educational needs and disabilities, and English as an additional language, if possible. This applies equally to the comparison group, if one is chosen – they should be as comparable with the intervention group across these demographic factors as possible. The key point is: it is important to think about any other external factors that could influence the results.

A pre-survey should be carried out before any activity has taken place, and at the same time across all groups, if at all possible. When the post-survey is carried out partly depends on the outcomes of interest – for example, tracking impact in three or six months' time – but it should be once the programme is complete.

There are a number of things to remember when administering a questionnaire:

- these kinds of questionnaires are often completed in hard copy, at desks – set aside some time in the classroom or during form to do so. However, electronic versions could be explored, which can save time on data entry and analysis;
- ensure that all students complete all questions if possible – be willing to help and facilitate if respondents have any questions;
- be sure to explain clearly how to answer each question, for example, demonstrating how to use a scale, so as to encourage the full range of responses – for younger students, consider going through the questionnaire question by question;
- if the results are to be anonymous, stress this to the students, as it could encourage them to be more open in their responses – also make plain that they should answer honestly, that there is no 'right' answer;
- as with a test, prevent students from copying or influencing each other's answers – the interest is in individual answers not the right response!
- make a note of contextual factors that might influence results, for example, if it is a particularly hot day and students are irritable, or they have just come out of a disruptive lesson. These factors will likely have an impact on the results of the questionnaire.

3.2.7 Analysis and interpretation

Once the questionnaires are completed and available in hard copy, it is time to convert this into data. Depending on the number of respondents, this could take some time – so be aware of the work involved!

The data is then analysed, to provide aggregate results across everyone who took part in the intervention.

To do so, follow these steps:

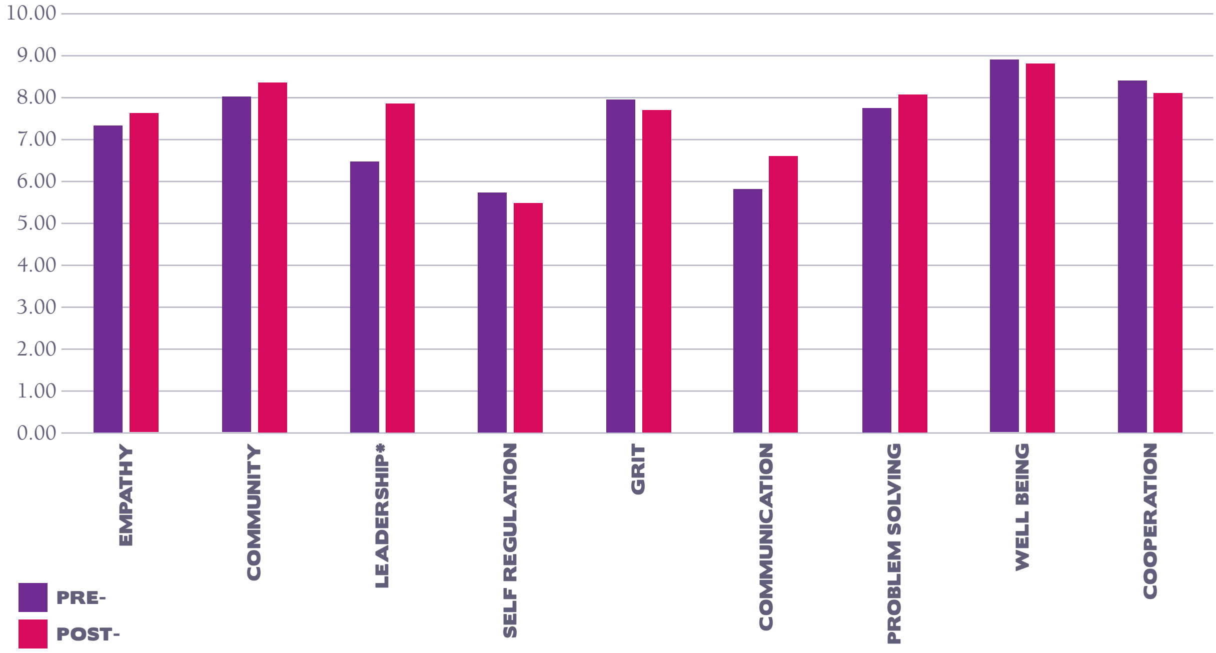
- calculate the pre-score for each character outcome for each individual by averaging across each question that makes up an outcome. In some cases this may be only one question, in which case, this is the score;
- do the same for the post-data;
- calculate the difference between these for each individual by subtracting the pre-score from the post-score. This will provide an indication of the change in each character strength for every individual. Caution is needed when interpreting the scores so that the level of 'effect', or size of difference, is also taken into account. If the score is positive there has been a reported improvement, if negative then a decrease;
- then average these 'differences' across all respondents – this will tell you if across the whole programme there has been a relative increase or decrease in the character outcomes you were interested in at the start. Note, however, that the success of the programme rests on the strength of this increase or decrease which is discussed below.

Results should be available at the end of this process – an average score for an outcome at the beginning and the end, and an idea of how relevant this is.

To give an example of what these results might look like, the key findings graph from the Character by Doing evaluation is provided below (see Figure 7). Do not be particularly alarmed if some character scores appear to have declined – evaluation for character development is not a perfect science and there is often noise amongst the signal. With smaller sample sizes, there will also be quite a lot of random error – meaning that the observed decline is not necessarily one that has occurred in reality. Testing for statistical significance gives an idea of what observed changes are unlikely to have happened as a result of random error, and this may be something to consider if a more advanced approach to your analysis is desired.

One way to further improve confidence that it was the intervention that influenced the observed change in character outcomes is to have a comparison group also complete the survey. If using a comparison group, what is needed is whether the change observed among intervention participants is any greater than among non-participants. To find this out put the comparison group data through steps 1-4 as above, and then calculate the 'difference of the differences' by comparing the change observed in intervention participants with that of the comparison group. An advanced approach may be to use Excel to undertake t-tests. Taking into account the size of the difference in scores between intervention participants and the comparison group could provide greater certainty that the intervention had the observed effect.

Figure 7: Example of the presentation of pre- and post-data



Drawing conclusions

Once the results are available and there is confidence in the calculations, they should be shared with colleagues; discuss what they mean for the chosen approach to character: perhaps focus a CPD or INSET session around them. The purpose of evaluation, particularly at the school-level, is less to 'prove' that a particular intervention works, but more to assess the strengths and weaknesses of an approach, and improve practice.

For this reason, alongside the impact questions it is particularly helpful to ask those involved - through questionnaires and interviews - what they enjoyed about the programme, what they thought was effective and less effective, and what they would change. This will keep driving improvement in practice, and have a varied evidence base on which to make judgements.

3.2.8 Summary: Surveys

- **Set out the research objectives:** this could be as simple as writing out the research question, or a logic model for the intervention could be developed. A clear idea of the intended outcomes should be in place.
- **Design the questionnaire:** draw on pre-existing examples to put together a questionnaire that assesses development in the outcomes of interest. Remember to design it with the audience in mind, and pilot it with a similar group, if possible.
- **Administer the questionnaire:** set aside some time in school to administer the questionnaire before and after the intervention, and compile the data.
- **Carry out the analysis:** calculate the average for pre- and post-scores, and review if there has been any change. If the sample size is large enough, test the results for statistical significance.
- **Share the results, and adapt accordingly:** once the findings are available they should be discussed with colleagues. Think about what they mean for future interventions and for individual character development.

References

Behavioural Insights Team (2016) *Evaluating Youth Social Action- Final Report*.
Available at http://www.behaviouralinsights.co.uk/publications/evaluating-youth-social-action-final-report/ (Accessed: 17 August 2016)

Bryman, A. (2012) *Social Research Methods*. Oxford: Oxford University Press

Character Lab (2015) *Character Growth Card*.
Available at https://cdn.characterlab.org/assets/Character-Growth-Card-8a9b995138cfd2572a42c2d34ba958e340211cde8ba2a1e80ab4488 7fb69c671.pdf (Accessed: 17 August 2016)

Demos (2016) *Character by Doing: Evaluation*.
Available at http://www.demos.co.uk/project/character-by-doing-pilot-evaluation/ (Accessed: 17 August 2016)

Education Endowment Foundation (2013) *DIY Evaluation Guide*.
Available at https://v1.educationendowmentfoundation.org.uk/uploads/pdf/EEF_DIY_Evaluation_Guide_2013.pdf (Accessed: 17 August 2016)

Youth in Mind (2016) *Strengths and Difficulties Questionnaire*.
Available at http://www.sdqinfo.com/ (Accessed: 17 August 2016)

3.3 MORAL DILEMMAS

Moral dilemmas are a method that can be used to demonstrate how individuals have come to a moral judgement. The sub-section considers how this method can be applied in a character education evaluation, along with taking a step-by-step approach to the development of moral dilemmas and their analysis.

Professor Steve Thoma,
The Jubilee Centre for Character and Virtues

3.3.1 Background

Moral dilemmas, or stories, that place the protagonist within a situation in which he/she must decide what is the right/just/fair/virtuous course of action have a long history within moral psychology. Popularized by Piaget (1932) and later Kohlberg (1981), dilemmas are seen as a powerful window into how the individual arrives at a moral judgement. Used in semi-structured interviews, dilemmas allow the experimenter to assess how the individual organises the defining features of the story and can identify which claims or issues are prioritised and similarly, excluded. For many years, dilemmas were the method of choice for gathering data and forming models about how moral concepts are understood and developed. As the field matured in the late 1990s and into the new century, an expanded set of methods and models of development were proposed and yet the use of dilemmas continues to be a featured component of many measurement systems.

Not surprisingly, given this research focus, applied researchers and teachers have adopted different strategies for bringing moral dilemmas into the classroom. A large and diverse literature exits about what does and does not work in promoting growth in moral judgements. In general, these reviews indicate that dilemma discussion is a useful approach to both assess students' characteristic ways in which moral judgements are made, and as the centrepiece of intervention strategies to promote growth. What follows focuses on more current uses of dilemmas in evaluating students' moral judgements and is informed by formal procedures used to establish reliable and valid measures of moral thinking in a variety of fields and age groups. Compared to earlier dilemma-based approaches using stories constructed by the experimenter, the procedures described here attempt to maximise participant engagement by using real-life dilemmas generated by members of the target population. Using students as collaborators in the measurement development process is hoped to enhance the relevance and plausibility of the target stories. Emphasising these two characteristics increases the participants' interest and also seems to better reflect characteristic ways in which people make real-life decisions. In the following text, the methods for developing these dilemma-based assessments are described and their potential for use in the classroom explained.

3.3.2 Developing appropriate dilemmas

Identifying the dilemmas used to focus participant interest and elicit moral thinking is central to the evaluation process. There are two main strategies to develop stories for our evaluations- group discussions and individualised assignments.

These strategies could be used in combination in a classroom, or either approach may be sufficient on its own. The test of whether to use one or both approaches is a judgement call based on when the teacher feels comfortable that the generated stories are both realistic, plausible and have generated a sufficient range of solutions for what the protagonist ought to do and why. More will be said about this point below but, in general, the story must be controversial and generate a sufficient range of options, both reasonable and problematic, to serve as an assessment of student application of the virtues. That is, if everyone sees the situation in the same way there is no variability to explore!

Group discussions

The process of generating stories typically begins with small groups of students considering real-life situations that highlight the virtue or concept under question. Specifically, when the focus is on school-aged children, the discussion can be prompted with a description of a virtue (or short list of virtues). The children are then asked for situations they have observed in their lives or within their peer group, where the target virtue was expressed. Each group can then be asked to highlight:

- Why they thought the virtue was present in the situation?
- What made the application of the virtue difficult?
- What they think most people would do in the situation and why?

During this process the facilitator typically circulates and listens for the quality of the discussion and, if it is flagging, asks for a synopsis of what has been discussed, raising questions to highlight aspects of the dilemma and hopefully stimulate the discussion. Each group is asked to identify a scribe who will write down the main features of the situation and also list the optimal solutions and the justifications for the listed actions. Ideally, at the end, each group is asked to present their work to the full group. This last step encourages effort and most importantly allows researchers to judge the reaction of others to the suggested situation and solutions. Note that the resulting stories rarely illuminate a single virtue and often what makes the story a dilemma is a clash between two virtues. For example, a common story generated in early adolescence is one in which honesty conflicts with loyalty. It seems that adolescents, particularly in the West, have great difficulty emphasising honesty and when doing so may negatively impact their friends.

Teachers can use the outcome of these exercises to provide a broad picture of how students apply virtue concepts, where they are clear in their understanding and where there are difficulties. This step may also suggest targets for subsequent interventions. This procedure (albeit with other target virtues), could be repeated at the end of the term in order to assess growth in virtue applications.

Individual assessments

Individual assessments complement the group work by estimating the range of responses within the age-range that may otherwise be masked by a group approach. For instance, it is always possible that adolescent group findings are influenced by the exaggerated input of the more popular or more outgoing individuals. To counter this possibility, participants are asked to respond individually to a series of semi-structured questions. Two different assessment options for this step have been tested. The most open-ended is an approach very similar to the group task. Typically this approach is used when the benefits of the group work are not available and a range of potential stories and judgements is quickly sampled. In this open-ended approach a definition of the target concept is provided and the student is asked to reflect on his/her own experience or the experiences of their friends and to present a situation that highlights the application of the virtue – particularly when it is difficult to do so. The student is then asked to reflect on what made the situation difficult. Finally the student is asked to list the possible outcomes that he/she would expect to observe that are admirable and those actions that might be more disappointing. This step is followed by a prompt asking the participant to explain the justification for these choices. Similar to the group responses, these individual student reactions are summarised to identify story themes and to generate lists of actions/justifications associated with each theme.

The open-ended version of the individual approach has the advantage of being the most sensitive to a variety of real-life interpretations of the target virtue or moral concept. However, the increased sensitivity comes at the cost of a higher frequency of weaker or non-compliant responses. It also suggests a premium is placed on written responses that may advantage higher ability students. To overcome some of these problems, particularly to reduce the number of weaker responses, the task is modified by using a ready-made dilemma. When used in sequence with the group approach these dilemmas may reflect one of the themes indicated in the group responses. Otherwise the facilitator may create stories that address common themes found in the literature or perhaps given in the curriculum. The task for students is similar to the series of questions described above, with some minor exceptions. If the dilemma is given to participants they are always asked whether the story is realistic and plausible. Greater focus is placed on the justifications and action choices by having the student list as many appropriate choices and justifications that come to mind as well as choices that might be expected that are less than admirable. These responses typically yield a wealth of information about what students think are acceptable and unacceptable choices and justifications. Additionally, an assessment of whether or not the students find the target story realistic and representative of their own life experiences can be made.

Acceptable and unacceptable choices

The interest in both acceptable and unacceptable choices arises from the observation that individuals often have more difficulty identifying unacceptable choices than the better ones. This surprising finding may be due to how students are instructed about the virtues with a focus placed on virtuous choices. Alternatively, the observed difference between acceptable and unacceptable choices could be simply because more acceptable behaviour is observed in everyday life. Whatever the cause of this disparity, it seems clear that acceptable and unacceptable choices provide unique non-redundant sources of information. Indeed, in professional populations the inability to detect bad choices and justifications on measures of moral thinking is associated with poor clinical performance. Similarly, adolescents who are acting out in school differ from their peers on how well they can identify inappropriate choices and justifications in moral situations and much less so on identifying appropriate options. Taken together, it is suggested that facilitators attend to both forms of choices and justifications in any evaluation of student performance and understanding of the virtue concepts.

Moving toward a more formal self-evaluation tool

Either of the above approaches for generating stories and choices yields a relatively detailed picture of the strengths and weaknesses in students' application of the virtues. For some facilitators the information provided in these earlier steps is sufficient for curriculum planning and interventions. However, if a more sophisticated evaluation is required, a more standardised measure can be developed. This next step mirrors the final process of developing a more advanced research tool, albeit using a more simplified and streamlined set of procedures.

Teachers should begin by identifying dilemmas generated by the students in one or more of the procedures suggested above. As emphasised earlier, stories that are particularly controversial, yield a range of action choices and justifications, and show evidence of misidentification of acceptable and unacceptable items should be selected. A story with these properties provides the greatest probability of identifying students who differ in their understanding of the virtue concept.

Once the stories have been identified, the next step is to construct roughly 8-12 items for both action choices (what the protagonist ought to do) and justifications (why he/she should pursue that course of action). These items represent distillations of the choices and justifications supplied by students and ideally are presented in a sentence or two. Limiting the detail given in each sentence minimises the possibility of over-interpreting the intent of the item. That is, simply by raising the issue the student must engage with the item to fully abstract its meaning and in so doing provides a more accurate indication of the students' understanding of the concept. For example, an action choice item for an honesty dilemma might read: 'Tom should not say anything to the teacher' rather than: 'Tom should not tell the teacher that his friends have an advanced copy of the exam because he….'. Similarly, a justification for the same story might read: 'I should never do anything to hurt my friends'. Note that these two examples do not represent an acceptable application of honesty as they fail to incorporate the purpose of the test, how the cheating might affect other students, among other limitations. As noted however, it is important to include items that the facilitator judges to be a poor reflection of the virtue because having both acceptable and unacceptable items allows one to estimate whether the student can fully apply the concept. One such example of a moral dilemma focused on courage has been noted below. Further information on this can be found in *Character Education in UK Schools* (Arthur, *et al*., 2015) along with Thoma *et al*. (2013).

Example of a moral dilemma about courage developed by the Centre

Madison, a year 10 student, was asked to represent her school in a gymnastics competition. A year 10 student had never been chosen to represent the school in gymnastics before so she was surprised at the opportunity. Madison realised this could have an incredible impact on her goal of becoming a professional athlete. However, after the first competition, she realised that she was not comfortable because the coach has been using the girls' good looks to get attention for the team from TV and newspapers. This included photographs of just the girls wearing their gymnastic outfits. She realised that using pictures of girls like this for publicity went against her own values and beliefs. Madison talked with her best friend about her concerns. He stated that if she missed all the publicity events then the coach would not take her seriously as a gymnast. Madison knew the importance of this opportunity. She had been told that if she did not take part in all team activities she would not be allowed to compete or develop her talent with the coach. However, she had always taken pride in her beliefs and in the examples that she had set, and if she let the coach use her good looks in this way, she would promote behaviours and values that she felt were very wrong.

Figure 8: Example of action and justification choices used in the moral dilemma about courage

ACTION CHOICES	JUSTIFICATION CHOICES
Madison should quit the team competitions after explaining to everyone her reasons.	It is just some photographs and TV shots.
Madison should compete but make sure everyone knows she does not agree with the photographs of just the girls for newspaper and TV publicity.	It is a great opportunity that could have a positive impact on her athletic career.
Madison should do what her parents think is best.	It would compromise her values and beliefs, which would make her unhappy.
Madison should talk to the coach and try to convince him to let her stay out of the photographs and media attention.	She would not be successful in gymnastics if she does not feel comfortable with the media.
Before she decides, Madison should think further about how her decision would impact her future.	She will be better off in the long run if she doesn't compromise her beliefs.
Madison should consider what would have the most impact on her friends.	She must believe that others know what is best for her.
Madison should see if changes could be made to the team publicity and photography arrangements.	Similar issues are likely to happen again so she has to figure out a way of dealing with them now.
A place on the team is an opportunity she can't pass up.	If she quits, the place on the team will be filled by someone else anyway.
	By competing in the competitions now, Madison may be in a stronger position to argue for a more appropriate situation with the publicity later.
	By standing up for her beliefs, Madison strengthens them through sacrifice.
	Beliefs come and go, but opportunities like this are rare.

3.3.3 Developing the measure

Once the dilemmas and items are selected, a standard format to create the measure is usually followed. This begins with an overall introduction to the task to ensure that a robust test can be developed.

This sub-section also includes an explanation on how to approach the various components of a moral dilemma. This includes the structure of moral dilemmas, how they are scored and how moral dilemmas can be applied to self-evaluation. The structure of a moral dilemma is noted below:

- Students are asked to read a story and are presented with a series of action choices which they are asked to rate in terms of their appropriateness on a five-point scale (From 'I strongly believe this is a good choice' to 'I strongly believe it is a bad choice' through a midpoint that indicates no endorsement; i.e. 'I am not sure').
- Students are then asked to rank the top three items that are considered the best choices from the set of options and then the three judged to be the worst choices. Students could also be asked to rank the top two items, particularly if the number of items in the section falls below eight.
- The ranking process is repeated for the list of justification items with only a slight modification to the wording used to describe the rating scale, changing 'good choice' to 'good reason' to describe the five-point rating scale.

In summary, the structure of the assessment follows a fixed pattern beginning with an introduction, the target dilemma, followed by action choice items both rating and ranking, and concluding with the justifications items, which are also rated and ranked. This sequence is repeated if a second or third dilemma is used. Typically for research purposes there are between four and seven dilemmas which can take an hour to complete. For the classroom however the number of dilemmas can be reduced and/or spread out over multiple class periods. Overall, however, the goal of the assessment is to provide information identifying which items the student feels are best choices and justifications, as well as best and worst justifications.

3.3.4 Scoring the assessment

Scoring of the measure is straightforward, as noted below:

- first the facilitator defines each of the choice and justification items as acceptable or unacceptable applications of the virtue. Items can be defined as 'neutral', indicating a choice or justification item that is not unacceptable but does not rise to the level of endorsement. These items ought to be infrequently used in the measure;
- the process of defining items can be very formal as a high level of precision is needed when identifying items as either acceptable or unacceptable. To achieve this level of objectivity, if possible, use an 'expert panel' – perhaps a group of teachers/students who have some expertise in the issues on which the dilemmas are based. During deliberation, the expert panel is asked whether it would be acceptable, unacceptable (or neutral) for a student to select the stated action or to provide the specific justification. The judging process is repeated for each action choice and justification until the group achieves consensus on the set of items. However achieved, the process ends when each item is assigned to one of the three categories: acceptable, unacceptable or neutral;
- to summarise a student's responses, the scoring system assesses how closely his/her good and bad ratings conform to the key. That is, students who rate good items as good and bad items as bad will achieve a high score. By contrast, students who endorse bad items as good and good items as bad obtain lower scores.

Mechanically there are a number of ways the choices can be aggregated to arrive at summary scores and it might be helpful to try different strategies. A possible approach is noted below:

- a weighting process can be used in which it is assumed that the first choice is closer to the student's preferred choice. Thus, when the top three best and worst choices are requested, the first rank is given three points, the second rank two points, and the last rank a point. This scoring process results in 24 points per story (six points each for the identification of the best and worst action choices and best and worst justifications). Misidentified items are also deducted. For example, if a student in the process of identifying the best action choices ranked acceptable items in the first and third ranks but ranked an unacceptable item in the second best rank, then the total for that section would be two (i.e., 3-2+1);
- if a neutral item is ranked a zero is inserted, indicating the item's status as neither inappropriate nor worthy of endorsement. Across sections this process results in an overall index of student performance expressed as a percentage (for example, the sum of the student's rankings over 24 for each dilemma);
- sub-scores can also be computed and can be very informative. Look beyond the total score (sum across all dilemmas and sub-sections) and assess: best vs. worst (combining the best action choices and justifications and comparing this sum to the equivalent score based on bad choices); and action choices vs. justifications (combining best and worst action choices and comparing this sum to best and worst justifications);
- in addition to the provision of an overall score, some additional calibrating information is provided which can be used to identify whether students are more adept at identifying good action or justification choices rather than the bad. Data are also provided as to whether students find it easier to identify good action choices or good justification choices. It may be of interest that it has been consistently found that justifications lag behind action choices in adolescent populations.

3.3.5 Applying the measure

There are many ways a measure of this type could be used. For someone just starting out, it might be prudent to begin with one of the group or individual approaches to gather student generated dilemmas and items as an exercise within the curriculum and then use this information to build a summary evaluation to be used at a later date. In subsequent years, the measure could then be administered to compare the effect of different modifications to the curriculum. Additionally, the measure can also be used in a pre- and post-design to better estimate change in student performance over the instructional period (e.g. is it uniform or do changes differ by gender, peer status, etc.). If multiple stories are constructed, students can complete them in sequence, following the instructional topics most relevant to the concept highlighted in the story.

3.3.6 Final advice

Using moral dilemmas in the classroom, both as exercises and assessments of character/moral education, provides a rich source of information about how students engage the targeted virtue concepts. If care is taken to ensure that the materials and tasks represent realistic and plausible moral dilemmas, the result is a comprehensive and nuanced assessment of the strengths and weaknesses of students' reasoning strategies. Information of this type can be important for understanding how educational practices influence student growth and how they may be strengthened to achieve maximum outcomes.

3.3.7 Summary: Moral dilemmas

- Moral dilemmas are used to evaluate students' moral judgements and, if formulated in a robust way, can be used to establish reliable and valid measures of moral thinking. They are centred on a protagonist who has a series of options of actions which the participant is then asked to rate according to their appropriateness. The example explored within this section also considers justification as to why the action is seen to be the most appropriate.
- Moral dilemmas should be plausible and relevant to participants. Consideration should be given to working with young people to generate the dilemmas. Dilemmas could be generated through group discussions or through individualised assignments.
- The results from moral dilemmas could be used in a number of different ways. The measure could be used both before and after an intervention has taken place to inform an estimate of its impact. Generating items with young people, and capturing their reflections on action choices and justifications could also provide some additional insights.

References

Arthur, J., Kristjánsson, K., Walker, D., Sanderse, W. and Jones, C. (2015) *Character Education in UK Schools, Research Report*. University of Birmingham, Jubilee Centre for Character and Virtues, [Online]. Available at http://www.jubileecentre.ac.uk/userfiles/jubileecentre/pdf/Research%20Reports/Character_Education_in_UK_Schools.pdf (Accessed: 17 August 2016)

Piaget, J. (1932) *The Moral Judgement of the Child*. London: Routledge

Kohlberg, L. (1981) *The Philosophy of Moral Development: Moral Stages and the Idea of Justice*. San Francisco; London: Harper and Row

Thoma, S. J., Derryberry, P. and Crowson, H. M. (2013) 'Describing and Testing an Intermediate Concept Measure of Adolescent Moral Thinking', *European Journal of Developmental Psychology* (special issue). 10, pp.239-252.

3.4 OBSERVATION

Practical advice on how observation of lessons can be incorporated into an evaluation of character education is provided by this sub-section. Key principles and a series of instruments which could be used to record and analyse the observation are also provided.

Matt Bawden,
Queen Elizabeth's Grammar School

CHARACTER TAUGHT

3.4

3.4.1 Introduction

Observation methods offer some particular advantages for teachers wanting to evaluate the character education provision of their school or a particular character education activity. Observation provides the opportunity to see how people interact and to see how they put their virtue knowledge and reasoning into practice.

Things to remember when undertaking an observation:

- observation is, by its very nature, formative rather than summative. It is important to see character education as a developing platform linked to a range of school improvement possibilities. As such, it is likely the observations will help to see the development of character education practice and lead towards, rather than show, an end point of best practice;
- it requires professional judgement. The more observation there is, the more likely it is to see virtue knowledge and reasoning being put into practice;
- when reviewing the observation there is a need for triangulation. Observations are powerful tools, but it is possible to miss things and to misunderstand. Using other methods, such as group interviews, to check what is found strengthens findings and adds weight when recorded in school plans and improvement documentation;
- student voice matters in an observation. How much an individual student understands about the virtues they are supposed to be acquiring during an activity is quite difficult to see. Asking them is one of the more precise means of finding this out.

3.4

3.4.2 Guidance on how to conduct an observation

It is possible to observe an activity, a lesson and more – even the manner in which the character virtues of the individuals within the school contribute to the building and maintenance of the whole-school ethos. The following points offer some guidance on how to conduct an observation of character education.

- Observing an activity such as queuing for the dining hall, corridor behaviour, or registration for lessons can show how character virtues are currently being used, and if revisited, how they are used after an intervention.
- Observing a lesson, whether one based on character education or on another curricular subject, will also provide fresh insight for the teacher, or his/her line manager.
- Observing such things repeatedly, and over time, will provide a detailed record to add to the evidence of spiritual, moral, social and cultural (SMSC) development and ethos.

Observing character education can be problematic without a clear plan and framework. It can appear as though there is far too much to see and consider. But it is really fairly simple. The steps are explored below.

Before the observation

Before starting an observation there are some things that should be done. The most important is to decide the evaluation question(s) it is intended to answer through observation. Examples might be:

- what is the influence of a particular character education intervention?
- how do students at this school demonstrate the virtue of compassion on a particular day?

Once the questions are set there needs to be a clear focus so the observer knows exactly what is to be observed, why, and how. The focus table (instrument 1) helps to make this clear. Filling this in now will save a lot of time later on.

Instrument 1: Focus table

Why observe?	List reasons for the observation
What is to be observed?	Outline the aspect of character development/education to be observed
How will the observation be made?	Summarise the method deployed to capture the observation
What will be the impact of this observation?	List effects of the observation and likely next steps

Preparing to observe: Experience the situation before you need to observe it

Wherever possible go and see what happens normally. Seeing what is normal will help the observer see what is not. It seems obvious, but failing to visit beforehand can make things trickier at the time of the observation, when so much else will be going on.

- Gather as much information as possible around the character virtues and the environment for the forthcoming activity to be observed.
- Think about who will be there and how they will impact on the character development activity.
- Consider how the location might have an impact on the development and acquisition of the virtues.
- Note what the students are learning in terms of the subject, whether they are already implicitly learning about the particular character virtues, and how their emotions are affecting their virtue knowledge and reasoning.
- Think about whether there are any ethical issues associated with doing your observation.

> Popping in to a maths lesson taught by the same teacher, with the same class, at the same time as the structured observation will take place the following week will really help to sort through some of the normal, or expected, behaviours to be expected the following week. Equally, watching the coach prepare his or her rugby team for the match will provide the foreknowledge to observe the team's resilience work in the match later that week.

The following list provides a series of questions you may want to consider when preparing an observation, and this list is also reinforced by the checklist included within instrument 2.

- How will the character work be observed? A basic decision needs to be made as to who will be observing. It does not have to be the lead person for the evaluation. Peers can observe, as can parents, governors and others. All they need is some clear advice so they know what to look for.
- What recording method will be used? Will writing a few notes be enough? Sometimes schools record the activity or lesson using either audio or film. It is probably best not to over-complicate matters; the easiest option is often the best.
- How will the review be carried out? Time needs to be set aside for looking back at the observation. This might involve both the deliverer and members of the observation group. Thinking now about when this will be can help keep people in the right frame of mind for meaningful feedback. Normally the sooner the review is completed the better, while things are still fresh.
- How will feedback be given? As with the review this needs time. Some of the best insights come out of a great dialogue with the person delivering the lesson, after all they were there before the observation, they know the students, and they will hopefully be able to see some of their own mistakes. These can then be used to develop the school plan, contribute to the self-evaluation framework, or on a smaller scale the individual teacher's continuing professional development.
- What will be done next? This could be almost anything from observing again, to re-writing the character development idea. Having some thoughts on this beforehand can be very useful. However, do not be tied to them. Excellent new avenues can open up at any time.

Instrument 2: Checklist

Clear focus		Observation method	
Recording method		Critical friendship/Peer review	
Feedback method		Next Step planning	

Deciding who will observe can be an important early step. If the activity is a part of a particularly technical exercise it may be worth considering pairing with someone who is a specialist in that subject. At other times a pastoral leader or tutor might be best suited for paired observation. This may be true when the activity sits within their remit during such things as form time or assemblies. When making these decisions, be aware that how such observations are explained to staff may also affect the outcomes. Where possible such observations should always be seen as non-threatening and meant to further enhance the school's ability to improve their character education. Observations may be tied to individual continuing professional development, but ideally should be viewed in relation to their role for whole-school improvement.

Can the finding be found again?

Can the behaviour be seen again? If it is a one-off it might not be that effective. Allow a set period of time in which it can be expected to see much of the character education behaviour, and stick to it. Martin and Bateson (1986) called this 'continuous recording'. This helps the observer see if the observed character education/development:

- recurs within the session;
- whether it does so at a sustained or higher level; and
- whether the impacts can be maintained.

A physical education session where the student needs to wait before commencing an activity might well be said to focus the mind on the virtue of patience. It is possible the teacher might make the student wait, then engage in the activity, and then explain the benefits of patience in that setting, as well as in other aspects of that student's life. All of this may be observed, and as such it can be said that character education has been seen to take place.

All we can observe, however, is the person experiencing the activity, and draw our conclusions on its effectiveness from this snapshot. If this activity is revisited a week or more later, improvements in the student's ability to wait, and to rationalise why waiting mattered, both in that activity and in wider life, might be found. Revisiting the rugby coach in the physical education lesson following the match might pose significant advances in the team's resilience; observing the same student entering the canteen each day for a week might show similar improvements in confidence.

Taking the example of the student in the physical education lesson, consider whether to record how the teacher sets up the activity, deals with late students, and more that might be found in any lesson in any school. It could be argued that how the teacher sets the activity and deals with lateness sets the tone for the character work, in which case it needs to be recorded. Such actions by the teacher have an implicit effect on the students. Nevertheless, a concern with the explicit intervention is more likely.

3.4.3 Completing the observation of an explicit character education activity or lesson

The questions below may help to focus an observation. A prompt sheet is also provided (instrument 3) and this is another resource that could be used to structure an observation.

1. Does the lead (teacher, teaching assistant, etc.) have high expectations and secure knowledge of character education?

2. Is the lesson well-structured regarding, and building upon, the principles of character education?

3. Is it clear how progress occurs in the knowledge, understanding and application of the chosen virtues?

4. How appropriate is the level of challenge in relation to character development?
   - The activity or lesson will be at the right level, neither too easy nor difficult for the students to grasp.
   - The subject content will be beneficial and not a hindrance, whatever the context of the lesson.
   - This means the subject level may need to be lower to allow a focus on the character work. There must be opportunities to reflect and learn when things do not go well, just as there needs to be opportunities to explore how to further embed good practice when it is observed.
   - Alternatively, it may also mean the subject content needs to be higher to force the need to develop character virtues or strategies.
   - Often the observation will include a focus on the use of dilemmas (for further information see sub-section 3.3). These can be very effective when seeing virtue understanding and application in the classroom. This can mean a reduction in lesson content to allow for deeper study.
   - If the observed activity is a repeat of previous work it still needs to be stimulating.

5. Is the attitude to learning about character positive, are students interested and engaged and enjoying the lesson?
   - Appropriate challenge creates a positive environment.
   - Students enjoy their learning and can see its purpose.
   - Sometime after the lesson students can remember the activity and its purpose.

6. If relevant, are support staff/students integrated into the method of learning around the chosen virtues?
   - Some students, who need support, are ably supported by others.
   - The students who struggle are well identified; these may not be the ones who normally struggle.

The following example demonstrates what an observation of a lesson could look like:

A science teacher who can list virtues may have good virtue knowledge, they may be able to link these to tasks where the students can independently consider each virtue, and they may plan times where the students can apply these virtues to their lives, to an area of science, or perhaps to their preparation for an assessment.

The science teacher might start with some key terms and then place these in the context of an experiment, before asking the students to use them in the writing-up of the method. Finally the teacher might ask the students to teach each other about their lesson's work.

To begin with, some hesitancy in the science lesson is to be expected, after all they are just beginning to learn, but as the lesson progresses things should change. Asking a student to explain how curiosity aids their science experiment will garner a reasonable answer from the outset, but asking it again once the student has performed the experiment is likely to gain a better response.

In the science lesson this might mean a student is curious, and very patient already, but may lack in other virtues such as resilience that are needed for the lesson to run smoothly. The observer spots whether the teacher might bolster the student's resilience, and offers timely advice.

Instrument 3: Lesson observation prompt sheet

The following instrument may be helpful to direct the focus of the observation.

PROMPT **CONSIDER HOW TRANSFERABLE THE ACTIVITY IS TO OTHER AREAS AND IF IT IS FIT FOR DELIVERY OF CHARACTER EDUCATION VIA THE QUESTIONS BELOW. RECORD INITIAL THOUGHTS IN THE SPACES PROVIDED BEFORE TRANSFERRING TO THE MAIN OBSERVATION FEEDBACK FORM (INSTRUMENT 4).**	**1 POOR** **2 SOME EVIDENCE** **3 GOOD** **4 OUTSTANDING**
1 Does the teacher have high expectations and secure knowledge of character education?	1-4
2 Is the lesson well-structured regarding, and building upon, the principles of character education?	1-4
3 Is it clear how progress occurs in knowledge, understanding and application of the chosen virtues?	1-4
4 How appropriate is the level of challenge in relation to character development?	1-4
5 Are the attitudes to learning about character positive, are students interested and engaged and enjoying the lesson?	1-4
6 If relevant, are support staff integrated into the method of learning around chosen virtues?	1-4
7 Is there an observable relation between the chosen virtues or method and attainment?	1-4
8 Is there an observable relation between the chosen virtues or method and behaviour?	1-4
9 Is there an observable relation between the chosen virtues or method and employability?	1-4
10 Is there an observable relation between the chosen virtues or method and school relationships?	1-4

Completing a character activity/lesson feedback form

In a more general observation, or even in the above context, a character activity/lesson feedback form (instrument 4) can be of value. It is easy to use, and over time a robust record of character development from across the school can be kept. This approach, from Kings Langley Secondary School, sets everything out simply and clearly, offering an easy way to record and set up the observation. By sharing the form ahead of time, those being observed are able to prepare and play a full role. Afterwards he/she can use it to plan improvements, and the observer can create an easy to use cumulative evidence file for the wider school.

Instrument 4: Character Lesson Feedback Form

<table>
<tr><td>NAME:</td><td></td><td>SUBJECT:</td><td></td><td>OBSERVER:</td><td></td></tr>
<tr><td>DATE:</td><td></td><td>CLASS:</td><td></td><td>VIRTUE:</td><td></td></tr>
<tr><td colspan="2">What we are looking for</td><td colspan="2">What we have seen</td><td colspan="2">What we are hoping to see</td></tr>
<tr><td colspan="2">How is the virtue incorporated into the objective?</td><td colspan="2"></td><td colspan="2">Either two objectives -
one for content and one for virtue
OR
combined objective</td></tr>
<tr><td colspan="2">Can the virtue be seen throughout the whole lesson? Is it explicit or linked to the lesson content or both?</td><td colspan="2"></td><td colspan="2">Looking for the virtue to be addressed at least once explicitly, but also to be linked to content throughout the whole lesson. Ideally the teacher will make reference to the virtue as part of conversations or teaching throughout the lesson.</td></tr>
<tr><td colspan="2">Is there the opportunity for the students to engage with the virtue? (discussion, debate, dilemma etc.)</td><td colspan="2"></td><td colspan="2">Look for at least one activity where students engage with the virtue explicitly. Ideally, other virtues will also be linked to the lesson content where appropriate.</td></tr>
<tr><td colspan="2">During the plenary/review of the lesson is the virtue reviewed
a) on its own?
b) by linking it with the subject content?</td><td colspan="2"></td><td colspan="2">Students have the opportunity to reflect on what they have learnt about the virtue explicitly and how it impacts on the content of the lesson.</td></tr>
</table>

3.4.4 Analysing the data from an observation

It is clearly important to use a form such as instrument 4, the character lesson feedback form, to draw out findings. The findings drawn out by using the questions within the various instruments can be followed-up and used to support the development of recommendations.

Following up the data

- Use the focus table (instrument 1) to highlight the original reasons for the observation. Following a character education observation there will be a lot of observed data, both qualitative and quantitative, to consider. Keeping the original purpose in mind will help to make the analysis sharp, and cut straight to the heart of the answers.
- Arrange evidence according to the original questions that were to be addressed, as noted in instrument 1. Organising the findings, perhaps with a couple of different highlighter pens or by setting up a table with the questions at the top of different columns, will show the amount of evidence/findings that support possible answers.
- Gather any additional evidence separately, as it will be interesting to see the questions these findings might address later. Using a third highlighter pen, or an extra column, will mean nothing is wasted from the observation. Character education is complex, and the observation may have addressed particular qualities but others were also apparent. If no note is made these will be quickly forgotten. If recorded and a future question involves that quality or those activities then the data can be quickly restored.
- Consider the impacts of the observation and adjust future planning and preparation accordingly. The observation may have met its aims. It may have fallen short, in which case why? Make sure future observations do not fall in to this trap. Or the observation may have almost accidentally exceeded expectations, again why? Build in this happy accident to future observation planning.
- Plan future observations at the same time as any follow-up interventions. This will ensure the evaluation instruments are fit for purpose from the outset, and allow the instruments to inform the intervention just as much as the intervention informs the appropriate instrument.
- Make careful recommendations. Think through the quality of the findings, just how much evidence there is for the answer to the original question, and make a judgment on the strength of each piece of evidence. Then act on it. This may mean trying another method to check the observation's validity (perhaps a group interview - for further information see sub-section 3.5 on interviews with students and teachers), or it may mean trying the same method in a different setting.
- Make the evidence a matter of record. Share the findings and ask for feedback.
- Celebrate the positives. Those involved have been engaged in important work; make sure they and others know its purpose and impact. More staff and students will then want to join in.
- Include others in the findings. Tell anyone who will listen, and even those who will not, about the answers to the questions. Teachers, students, leadership, parents, governors and the community can all celebrate or help. The observation may have been uncomfortable for the activity leader, including him or her may change this feeling and encourage others to participate in the future.

3.4

3.4.5 Summary: Observation

- Observations of lessons, and other activity in the school, should be used in combination with other methods and provide a formative rather than summative evaluation of character education provision.
- Asking students about their view on the lesson, for example, what they understand and what they have learnt from the lesson may help to generate insights with greater depth.
- Observations should be guided by the research question and the intended results. Having this as a reminder may help to focus the notes.
- Consider how frequently what is observed can be seen again and if there are any differences in how this behaviour is demonstrated.

CHARACTER TAUGHT

3.4

References

Martin, P. and Bateson, P. (1986) *Measuring Behaviour: An Introductory Guide*. Cambridge: Cambridge University Press

3.5 INTERVIEWS

Interviews are a flexible approach that can be adapted for use with both adults and students, both in a group setting and on a one-to-one basis. The following sub-section goes through the different types of interviews that could be used, along with some practical advice on how they can be administered.

Matt Bawden,
Queen Elizabeth's Grammar School and
The Jubilee Centre for Character and Virtues

3.5.1 Introduction

Interviews are a flexible approach to gather a range and depth of information focused on answering a research question. They are also adaptable and, depending on the type of interview that is selected, allow the researcher to explore new ideas generated by the participant and to probe for more information (Bell, 2005).

An interview can be designed to gather in-depth insights, or more structured to provide quantitative data (Bryman, 2004). In addition, an interview can be done on a one-to-one basis or within a group setting and can be designed to be suitable for a range of ages, including staff and students.

This sub-section is structured in the following way:

- the different types of interview;
- things to think about when planning and conducting an interview;
- group interviews with students;
- semi-structured interviews with teachers.

3.5

CHARACTER TAUGHT

3.5

3.5.2 The different types of interview

Whilst the foundations of an interview are the same, i.e. questions are asked and responses are captured, there is a variety of ways in which an interview can be structured which can affect its design and the data collected.

What is a group interview?

A group interview is a research method which brings together a collection of people to discuss the research topic led by a moderator. There will be a series of questions or topics listed in a discussion guide developed before the group interview. This discussion guide is used to help to explore what people think about an issue and how they have come to this opinion (Bell, 2005). Consequently, it is a flexible approach which can allow participants to discuss the research question in a fluid way. Bryman (2004, p.352) notes that moderator involvement should not be 'intrusive or structured' (p.352), and that there should be scope to explore unforeseen topics, while still ensuring a focus on the research question.

Often, a group interview work best when covering a topic which is familiar to the participants. Evaluating character education from the perspective of the students lends itself to this method, as each participant will have a shared experience of attending a particular school and is likely to have views about their experience that could help to generate discussion across the group.

Group interviews and focus groups: a note on terms

There are a variety of views as to whether a group interview and a focus group should be considered to be the same method. Bryman (2004) notes that two distinctions are often made between group interviews and focus groups:

- focus groups are often based on a specific theme or topic, whereas there is the tendency for group interviews to have a wider span of interest;
- a focus group moderator will be interested in looking at how participants discuss particular topics within a group.

For ease of reference, this sub-section will use the term group interview. It is recognised however that 'meaning', or the shared interpretation of what a participant says, is constructed jointly between the moderator and participants. Considering how participants interact, including the level of agreement and disagreement between them, should also be of interest.

Evaluation of a school's character education provision means that it is likely that a broad array of topics will be raised by participants. It will also be of interest to look at the ways that students interact while discussing the topics, including looking at the level of agreement and disagreement (Bryman, 2004). This takes the approach that people's views are developed with reference to other people and that the interaction between participants is an important part of the approach. In this way group interviewing is seen to be different to individual interviewing (Gibson, 2007; Barbour and Schostak, 2005).

Structured interviews

A structured interview starts out with clearly defined questions developed in advance, restricting the researcher's freedom to stray off topic and explore interesting comments which the participant may contribute. This interview format however can be easier to analyse (Cohen *et al*., 2007). This is explored in more detail when looking at interviews with students and in the final part of this sub-section, which considers the analysis of interviews.

Semi-structured interviews

A semi-structured interview starts with a series of questions which the interview is intended to cover, listed in the discussion guide. The purpose of a discussion guide is to identify particular topics which are used to structure the interview (Bell, 2005). Nevertheless, the researcher has the freedom to ask additional questions and to explore additional areas of interest as they are identified. This type of interview will be considered in more detail when looking at interviews with teachers.

Unstructured interviews

Unstructured interviews are focused on a topic rather than being controlled through a framework of questions, also known as a discussion guide. This approach can allow new topics and insights to emerge which may not previously have been considered by the researcher and such insights can be useful in the preliminary stages of research. This type of interview can be difficult to manage and very time consuming to analyse. Consequently, this sub-section does not include further information on this type of interview.

3.5

3.5.3 Things to think about when planning and conducting interviews

The below list considers the things to consider when planning and conducting interviews.

Keeping a record

When undertaking an interview it is important to keep a record of everything said as this will support analysis. It is possible to record things by hand as notes, with a voice recorder, by video, or in combination, and each approach has its advantages and disadvantages.

- **Recording with notes:** Advantage – the researcher will only note key points. Disadvantage – the researcher may seem distracted to the students and may fail to note something earlier which becomes more significant later.
- **Recording with a voice recorder:** Advantage – cheap to organise, no comments are missed, and easy to set up by placing the recorder in the middle of the group. Disadvantage – can be a little intimidating until the group forgets about the recorder.
- **Recording with video:** Advantage – it is easy to look back on key events and findings can be kept as an edited video. Disadvantage – expensive and possibly too intimidating, with the student group taking longer to forget about the video recorder compared to the voice recorder.

Recording with notes and a voice recorder

In most cases this is the best way forward. Digital voice recorders are relatively inexpensive, recordings can be downloaded to a computer and record almost everything the researcher might need. There may be ethical issues that need to be considered if an interview is being recorded and students should give their informed consent that they are happy for the interview to be recorded. Within a group interview, all participants will be aware of others' responses, so the ground rules of the session should establish that respect should be shown to participants' answers outside of the session.

The basic procedure for recording a group interview could be as follows:

- inform the group and ensure they are comfortable;
- set the recorder going;
- state the date and time;
- ask each student to state their first name (and any other necessary information);
- start the questions;
- use a notepad to add any details the recorder will not catch;
- conclude by stating the date and time again; and
- if there is time, listen to the recording soon after to make sure all is clear.

Keeping a record

Assuming the researcher has not made a video record it is now important to transfer the notes to a record form (instrument 1). This enables the researcher to keep findings in a common format. It is general and yet precise enough to allow for future comparisons between group interviews and across interventions. By using a record form the researcher is also forced to distil any information down to key points. The researcher can, of course, still keep any voice recordings in a folder on a computer.

Instrument 1: Record form

PRE-INTERVENTION		POST-INTERVENTION	
RESEARCHER		DATE	
NUMBER OF PARTICIPANTS/ YEARS		PROJECT TITLE	
Qualities explored			
Notes on knowledge			
Notes on understanding			
Notes on application			
Areas for follow-up			

Analysing the record

Transcribing of interviews should be considered. The amount of time it takes to prepare a full transcription should not be underestimated, so it may be preferable to make notes from the audio file.

Rubin and Rubin (2012) provide guidance for the analysis of qualitative data. One practical piece of advice that may be helpful is to create a notable quote file which is used to jot down interesting quotations from the interviews. This file could be used to identify concepts and themes that could be methodically explored.

Coding is a key part of analysing an interview transcription. The following points explain how to code a transcription.

* Coding refers to the attribution of a word of phrase to a section of the transcription. It is possible for multiple codes to be ascribed to the same section of the transcription. Rubin and Rubin (2012, p.194) suggests that the following could be used to structure the coding of transcriptions: background information that may have influenced the topic being discussed; examples; concepts (ideas that relate to the subject discussed); and themes (defined as 'summary statements, causal explanations or conclusions'). Looking for variation in what has been said should help to capture a greater depth of comment.

- While there is software available to support coding, it may be preferable to use colour coding, through using the highlighter function or the comment function in Word to code transcriptions. Alternatively, coding can be done on a hard copy of the transcription.
- Once the transcription has been coded, these codes can be organised and compared with one another. The researcher can then attempt to identify explanations for what has been noted. It will also be important to consider the extent to which these are situation specific issues and whether or not the lessons learnt may be applicable in other settings.
- This handbook continually refers to the importance of triangulation when undertaking self-evaluation of character education and cross-referencing findings from interviews with other research methods will add further depth to the conclusions.

3.5.4 Group interviews with students

Interviews with students are a popular way to evaluate the character education provision of schools. They can be used to get a group of students' opinion on the school ethos and culture or to feedback on a particular character building activity.

General guidance for group interviews

There are several things to consider when interviewing students:

- what questions is the evaluation designed to answer? These will guide what questions are asked in the interview, who is asked and when they are asked;
- what questions are worth asking, and how will they be phrased? Well-phrased questions can open up conversations and can lead to deep responses. Trial them first with one or two students, and ask them for their opinions via a range of planned open questions;
- how should the answers be recorded? Writing notes can make the researcher seem distracted; using a voice recorder can be intimidating; video can be off-putting. Think it through carefully. There may also be ethical issues around recording;
- how will the answers be reviewed to develop findings? If the answers were recorded in note form they need quick editing before the observer forgets the context; if the answers were recorded on a voice recorder they need early analysis so the observer remembers who said what. The use of a video recording means it is less of a concern if analysis is delayed as everything from the interview is captured.

When could a group interview be used?

A group interview can be used to help gather data to inform the evaluation of both taught and caught approaches through gathering students' perceptions of their experience of character education.

Group interviews are likely to take place across peer groups and it may be that some participants will be more reticent if discussions touch on sensitive topics. It may be helpful to consider beforehand what topics are likely to come up in a group interview and to consider how discussions can be managed. The information on question types later in the sub-section may help to inform an approach to managing these discussions.

The role of teaching staff as moderators

- Teachers should be aware of their dual role of being both a moderator and a teacher and the potential impact this may have on the group interview.
- The moderator also needs to support the development of group dynamics between participants and also ask questions that will clarify participants' views where necessary (Horner, 2000). Both verbal and non-verbal feedback can be used to create a non-threatening environment and ensure participants feel comfortable to contribute their opinions (Horner, 2000; Rubin and Rubin, 2012).
- A power imbalance is also likely to affect interaction between the adult researcher and student participants (Horner, 2000). Participants may look to moderators for guidance and reassurance before answering questions and so moderators will need to be careful not to lead participants' answers (Gibson, 2012). Group interviews are focused on the interactions between participants however and this can reduce the power and the influence of the researcher (Heath *et al.*, 2009).

- ✳ Setting some ground rules at the start of the session may help navigate between the roles of being a member of the school staff and a moderator. This could establish that students should feel free to express their opinion and that they will not get in trouble if they express thoughts which are unfavourable to the school or to particular subjects, along with the expectation that serious responses should be provided to the questions (Gibson, 2012).

Group interview composition

When planning a group interview the following should be considered:

Age range

A teacher is in a good position to decide on an appropriate age range between students for the group interview. Hennessy and Heary (2005) suggest that an age range of 1-2 years will ensure that everyone is able to contribute to discussions at a similar level.

Mixed-ability group

Consider how a mixed-ability group could contribute to gathering a range of experiences and provide further depth of information.

Gender

When developing group interviews consider the extent to which the research question may be best answered by conducting interviews with single or mixed gender groups (Kennedy, *et al.*, 2001). As part of the moderator's role in responding to group dynamics, there may be a need to consider whether there are any gender differences in the frequency and depth of responses. The moderator should seek to ensure that both genders are given the opportunity to contribute.

Friendship groups

It may be easier for participants who know one another to open up and discuss topics with greater fluency (Barbour and Schostak, 2005), but note that there may be 'past histories and pecking orders at play' (Barbour and Schostak, 2005, p.43).

Size of the group

The age of the participants is likely to have an impact on the size of the group to be interviewed. Larger groups are likely to work better with older participants (Gibson, 2007). In contrast, Bryman (2004) notes that larger groups may not be suitable when participants are not familiar with the research topic, whereas smaller groups may be better for exploring topics in greater depth. Larger groups may be more difficult to control however, whereas in smaller groups it can feel like a series of simultaneous interviews are taking place. Consequently, it is recommended that a group interview is comprised of 5-8 people (Hennessy and Heary, 2005).

Keeping everyone talking

- ✳ The questions posed during a group interview will be influenced by the research question and the moderator will also need to respond to comments brought up during the interview. To maintain and develop discussion however, asking 'what' and 'how' questions that focus on experience may be more appropriate for young people's shorter attention spans, in contrast to more abstract 'why' questions (Rubin and Rubin, 2012).

- The balance between open and closed questions will also require additional thought. While closed questions may hamper discussion, it has been suggested that opening the group interview with questions that require short responses may warm-up participants and help them feel more comfortable (Gibson, 2012).
- Strategies to deal with silence may also need to be developed. One approach may be to ask participants to explain what the question means in their own terms. It may also be helpful to summarise discussions and to ask for confirmation whether this is a correct interpretation, along with asking participants to clarify what has been said after giving a prompt (Gibson, 2012).

Using a silent debate to keep the interview on track

Silent debates as an approach can encourage participants to develop their responses by asking them to write down their thoughts.

HOW TO HOLD A SILENT DEBATE

- Print research questions on large pieces of paper, and number them.
- Sit participants in groups of three.
- Place a large piece of paper in front of the participants.
- Give each a unique coloured pen.
- Ask the participants to sign their names on the reverse.
- Tell the participants to answer the question using their pen instead of speaking.
- They can comment on each other's answers or make new points.
- Set a time limit.
- Move the sheets round the groups and repeat.

Once the sheets have been completed they can be referred to during the remaining group interview time. Each individual will be identifiable by the pen used and through their signature on the back of the sheets. As a side note, it is worth numbering these sheets, and using these numbers when posing questions, to help when reviewing the recording of the group interview later.

Question types

Each group interview will be different according to the topics being discussed, but the list of question types below, reproduced from Rubin and Rubin (2012, p.116-120), may help to structure discussions.

MAIN QUESTIONS

- These are the 'scaffolding' of the interview and should be directed at answering the research question. They should be prepared in advance and structured in a way that participants will understand, and should also reflect previous points of discussion as this can demonstrate active listening.
- Types of main questions can include (Rubin and Rubin, 2012, p.136-147):

Tour questions: These are a broad description of actions and events, for example, asking participants to consider the average school day or lesson.

Experiences: These could ask participants to go through a particular experience, for example their involvement in extra-curricular activities.

Hypothetical examples: These questions pose a particular scenario and ask participants to reflect on their experiences. For example, asking about what an ideal school would be like.

Meaning of a key term: This asks participants to define a key term and can be a way of identifying the meaning they place on a particular topic. For example, it may be interesting to ask participants what they think is meant by 'virtue'.

FOLLOW-UP

- These questions respond to what the participants have already said and can be used to gain further clarification, detail and richness.

PROBES

- Probes help to gather further elaboration of answers that have already been given and can help the moderator to manage the group interview.
- Asking for clarification about the order of events can help to explore how these events are linked together, while asking for confirmation can ensure that what the participants have said has been understood correctly. Non-verbal probes can also be used to prompt the participant to continue, which could include using pauses to allow the participant to finish what they are saying, along with employing body language to invite further comments.

Further examples of questions which could be used to evaluate taught character approaches with students are listed below.

Evaluating character taught

1. What can you remember about <insert programme name>?

What can you say about each part of the programme? N.B. This would need to be targeted to the particular programme/project

2. Was there anything you enjoyed about the programme?

If you were telling a younger student about the project, what would you tell them?

3. Was there anything about the programme that you didn't enjoy?

If we did the class again, is there anything that you would do differently?

4. If your best friend asked you what a virtue is, how would you describe it?

Do you think that virtues are important? If so, why?

What virtues do you think are the most important? Why is this?

5. Do you think the programme helped you to learn about the different virtues?

What did you learn about the different virtues?

Did learning about the virtues help you change how you think about them?

How do you think you'll apply the virtues <INSERT> to your everyday life?

Can you list some examples of virtues?

3.5

3.5.5 Structured interviews with students

Using quantitative questioning (also see sub-section 3.2 Surveys)

- Structured interviews with students can also be used to gather a large amount of information in a relatively short amount of time. An evaluation of character education provision could combine both qualitative group interviews and quantitative structured interview approaches. Indeed, this may be one way of exploring the validity of findings. There are similarities between the guidance on the use of quantitative questioning and sub-section 3.2 on the use of surveys with students, and some of the advice could be applied when conducting structured interviews. Structured interviews may form the starting point for the discussions in the group interview. The pre- and post-intervention quantitative questioning tool (instrument 2- below) offers a few basic ideas as to the sort of questions that could be asked.
- There are difficulties in measuring character and virtue, as discussed throughout the handbook, along with the need to triangulate methods. A mixed approach, making the most of quantitative opportunities to illuminate qualitative responses is likely to generate a greater depth of analysis for school plans or continuing professional development.

Instrument 2: Pre- and post-intervention quantitative questioning tool

How far do you agree with the following statements?	Circle a number for each statement: 1 Strongly Disagree 5 Strongly Agree	(For use post-intervention) For each statement please state whether or not the activities/lessons helped you to do this. Please explain why this is the case.
I know how to act for the right reasons in different situations in school.	1 2 3 4 5	
I know how to act for the right reasons in different situations out of school.	1 2 3 4 5	
When I face a challenging situation, I am able to identify the right things to do.	1 2 3 4 5	
I understand the meaning of virtues	1 2 3 4 5	

3.5.6 Semi-structured interviews with teachers

Undertaking one-to-one semi-structured interviews with teachers is a way of gathering teachers' thoughts and evaluating character education.

Semi-structured interviews allow the researcher to explore participants' comments in detail. The below list summarises some additional areas to consider if conducting interviews with teachers.

- There is the opportunity to consider interviewing a range of school staff.
- Participants may be interviewed by their colleagues and the school hierarchy may prevent some participants from contributing freely.
- Interviewing teachers may allow a process evaluation to be carried out. This may add greater depth to analysis of the impact of a school's approach to character education (EEF, 2013). Developing a logic model (as referenced in sub-section 3.2) could provide a structure to generate questions on the implementation of character education.

Example discussion guides for evaluating character education with teachers

The example below is a discussion guide that could be applied to support discussions of both taught and caught approaches to character education with teaching staff. This is a general discussion guide that could be targeted to ask about different aspects of character education.

Introduction

1. How is character education implemented within the school?
(Prompts: Could you describe the school ethos? / Could you talk through a typical school day and how character education is embedded?)

Character education: caught

2. Do you feel the school has a clear mission to develop character?
If yes, how is this communicated? How is it defined?

3. Does the school have a clear list of core values or virtues?
If so, which of these would you prioritise as being the most important when teaching young people?

4. Have any parts of the school ethos contributed to students' character development?
If so, what? Which have had the most impact?

5. Is there any particular focus placed on ensuring quality student/teacher relationships?
If yes, what does this involve?

6. Have you had any training about character education?
If so, what training activities were the most memorable/successful?

7. What impact, if any, has the school ethos had? How has this been demonstrated?
(Prompts: What impact, if any, has the school ethos had on staff/students?)

3.5

Character education: taught

8. Has character education been embedded within the curriculum at the school? If so, how? What factors influenced the decision to do this?

9. Has character education been embedded within whole-school activities? If so, how? What factors influenced the decision to do this?

10. What impact, if any, have 'taught' approaches to character education had on students? How has this impact been demonstrated?

11. Does the school's approach to extra-curricular activities incorporate taught character education? If so, how do extra-curricular activities reflect this?

N.B. Asking participants to consider barriers and enablers may help to inform an evaluation of how character education has been implemented, which may help to identify whether or not the process of implementation has had an effect on the perceived impact of character education.

Barriers and enablers

12. What advice would you give to schools thinking about developing their approach to character education?

13. Was there anything that helped enable the implementation of character education at your school? If so, what?

14. Did anything inhibit the implementation of character education? If so, what? Did the school overcome these barriers? If so, what helped?

3.5.7 Summary: Interviews with students and teachers

- Semi-structured interviews offer the researcher the opportunity to probe the responses of participants to generate a greater depth of insight. Thinking about how questions can be structured and how other activities can be incorporated within the interview can help to develop a strategy of what to do in case discussions begin to dry up.
- Where the roles of researcher and member of the school staff are combined, it may affect the answers participants provide. Setting out clear ground rules at the start of interviews may help to limit the impact of this.
- Recording interviews will help with analysis, but the ethical considerations of doing so should be assessed.
- When analysing the data generated by interviews, it is important to consider whether these findings are situation specific or if they can be applied to other settings.

References

Barbour, R. and Schostak, J. (2005) '*Interviewing and Focus Groups*' in Somekh, B. and Lewin, C. (ed.) *Research Methods in the Social Sciences*. London: Sage, pp.41-49

Bell, J. (2005) *Doing your Research Project: A Guide for First-Time Researchers in Education, Health and Social Science*. Berkshire: Open University Press

Bryman, A. (2004) *Social Research Methods.* Oxford, Oxford University Press

Cohen, L., Manion, L. and Morrison K. (2007) *Research Methods in Education.* Oxon: Routledge

Education Endowment Fund (2013) *Evaluation DIY Handbook*. Available at: https://v1.educationendowmentfoundation.org.uk/uploads/pdf/EEF_DIY_Evaluation_Guide_(2013).pdf (Accessed: 17 August 2016)

Gibson, F. (2007) 'Conducting Focus Groups with Children and Young People: Strategies for Success', *Journal of Research in Nursing*, 12 (5), pp.473-483

Gibson, J. (2012) 'Interviews and Focus Groups with Children: Methods that Match Children's Developing Competencies', *Journal of Family Theory and Review*, 4, pp.148-159

Heath, S., Brooks, R., Cleaver, E. and Ireland, E. (2009) *Researching Young People's Lives*. London: Sage

Hennessy, E. and Heary, C. (2005) *'Exploring Children's Views through Focus Groups'* in Greene, S. and Hogan, D. (ed.) R*esearching Children's Experience: Approaches and Method*. London: Sage, pp.236-252

Horner, S. (2000) 'Using Focus Groups Methods with Middle School Children', *Research in Nursing and Health*, 23, pp.510-517

Kennedy, C., Kools, S. and Krueger, R. (2001) 'Methodological Considerations in Children's Focus Groups', *Nursing Research*, 50, (3), pp.184-187

Rubin, H. and Rubin, I. (2012) *Qualitative Interviewing: The Art of Hearing Data*, *(third edition).* London, Sage

3.6 EVALUATING CHARACTER EDUCATION LESSONS

3.6

The nature of character education means that there are some clear boundaries that affect how a schools' provision should be evaluated. This sub-section considers these boundaries and argues against the individual assessment and use of grading within character education.

Ian Morris, Wellington College

The sub-section concludes with a number of tools which could be used to support reflection on students' progression within character education.

CHARACTER TAUGHT

3.6

3.6.1 Introduction

Evaluation tools should give accurate feedback on learning progression, and the quality of that feedback should motivate students to continue to learn and make additional progress. Character education is intrinsically linked to the core identity of learners and this ought to place some constraints on the way in which it is assessed. These constraints are set out below:

- **No grades for character education.** There is well documented research which shows that if feedback to children comprises a formative comment and a summative grade, they look at the grade and ignore the comment. There are also some questions raised by the prospect of character grading and the interpretation of a numerical grade. In learning to acquire the habits of a good, virtuous life, an individual needs specific feedback on how they handle their emotions, the reasons they give and take for their actions and the way their emotions and reasons translate into action so that they can learn. Numerical grades do not give specific feedback, either summative or formative.
- **Social comparison.** Number grades could also mean it is more likely for students to compare themselves against others.
- **The interior life of the child.** Character education raises the challenge of working out what is fair game for assessment. In *The Dangerous Rise of Therapeutic Education* (2009), Kathryn Ecclestone and Dennis Hayes raise grave concerns about assuming the emotional vulnerability of all children and seeing the role of education as being principally about 'kid-fixing'. Where education is seen as *therapeutic*, Eccleston and Hayes suggest that there is a tendency for adults to go digging around in the interior lives of children, looking for problems which they can fix. This raises a boundary issue for character education which teachers need to be wary of. Maths teachers are on reasonable territory when interrogating a student's approach to solving mathematical problems in order to help them solve those problems better. When it comes to character development, there is the potential for asking much more probing questions about the nature of the self which teachers should be very wary of asking.

  Teachers ought to be clear about the distinction between what is visible (i.e. behaviour) and what is interior (i.e. thoughts and feelings) and not assume that the latter is open for interrogation. If a child wishes to keep their interior life private, that is to be fully respected.

Promoting self-reflection in students (for further information see section 4 of this handbook) is one way in which the boundary challenge for character education can be resolved. Self-reflection supports the student to consider their actions and the thoughts and feelings that influenced their behaviour and actions. Such reflections are then intended to influence the student's behaviour and their self-perception of their actions. The teacher can pass comment on the behaviour and the student maintains control of their inner life through reflecting on their behaviour and their teacher's interpretation.

3.6.2 Using the caterpillar process for lesson evaluation

The Centre's *Teaching Character through Subjects* (Harrison, *et al*.,2016) provides materials for character education and describes a 5-stage process of virtue development called the caterpillar process, after Eric Carle's popular children's book *The Very Hungry Caterpillar.* The five stages are as follows:

1. **Stop.** This involves the ability to either pause before making a decision, or to pause afterwards to reflect upon decisions taken. It is based upon the skills of emotional regulation that enables an individual to pause before leaping into a situation.
2. **Notice.** This involves gathering more information about situations rather than just going with the first thoughts, finding out more in an attempt to see situations as they are, rather than as desired.
3. **Look.** This involves observing our own emotions and the emotions of others. The emotions give information about what we and others perceive and these perceptions are not always appropriate. The *middle way* involves individuals feeling the right things, at the right time, in the right way, towards the right people, for the right reasons.
4. **Listen.** This involves the giving and taking of reasons for the things that are done and the feelings that are felt. Aristotle encourages the education of emotions with reason and the education of reason with emotions to slowly refine responses. This stage also encourages listening to the knowledge of the virtues and the practical application of it to situations.
5. **Caterpillar.** In *The Very Hungry Caterpillar*, the colours of the food the caterpillar eats end up on the wings of the butterfly he becomes. In the same way, what an individual thinks, feels, says and does end up 'colouring' the person that he/she is becoming. This final step involves reflecting upon responses across the previous four stages, with a view to changing responses that are lacking in virtue.

This reflective process can be applied to any learning done across the four domains of virtue (moral, civic, intellectual, performance), as any form of learning involves pausing, gathering information, feeling, thinking and reflecting to improve future performance. In the early stages of understanding this process, time will need to be taken to deliberately develop each aspect. For example, 'pausing' can be developed using mindfulness techniques and 'looking' will involve developing an understanding of how emotions manifest themselves in the body through feelings, facial expressions and posture. Very young children may not yet have learned how to identify and read these signals.

The caterpillar process can in turn be used to create a character development grid, the primary aim of which is to assist students in becoming aware of their character learning, not to provide assessment evidence for its own sake. It is important to reiterate that a character development grid should exist to help children create a mental landscape of learning within a discipline so that they can make progress, not so that they can be ticked off or scored. It is also important to stress that the acquisition of virtue is a life-long project. Some children experience levels of neglect in their early years which will make it difficult for them to embark upon moral learning until their emotional development and attachment styles are repaired. In school-age children, particularly of primary age, it is unlikely to see a child going beyond doing what is right through gritted teeth and being swayed by pursuing pleasure rather than doing what is right or good. This is in no way to denigrate the moral ability of students, who often surprise and inspire us with their insight and wisdom. Virtues are stable states of character however and few students (and surprisingly few adults) will be at the stage where managing emotions, being in command of desires and acting for the good are habitual, rather than exceptional. Moral development takes a long time. It is also not necessarily linear; just as good habits develop over time, bad habits can appear and undermine virtue learning that has taken place.

3.6

3.6.3 Self-assessment using a character development grid

An additional area which may prove helpful to consider is the provision of feedback on character development. This feedback does not have to be limited to being between teacher and student. Self and peer assessment can be a very powerful way of developing understanding of how well virtue is being acquired.

Students can be invited to write reflections on how successful they have been with specific elements of acquiring virtue, especially where they have encountered events which have challenged them. Further examples of approaches to self-reflection can also be found in section 4.

The grid below provides a framework which could be used by students to structure both their own self-reflection and that of their peers. In trying to establish a community of virtue learners, students can use the grid to give feedback to each other on the progress they are making, again especially when they have encountered challenges.

It may be that some students have achieved different stages at each phase of the caterpillar process and so this framework should be used as a resource to spark discussion when providing formative feedback, with a view to identify areas for development and to consider what action would support progression to the next stage of the caterpillar process.

Figure 9: The caterpillar process: Stages 1-3

	STAGE 1	STAGE 2	STAGE 3
STOP: pausing before moral choices are made.	✳ I did not pause to assess the situation. I was carried into action by my emotions or my desires. ✳ I was only interested in avoiding pain or pursuing pleasure.	✳ I found it difficult to pause to assess the situation. ✳ I was too easily influenced by avoiding pain or pursuing pleasure.	✳ I paused to assess the situation. ✳ I could resist the desire to avoid pain, but could not resist the desire to pursue pleasure.
NOTICE: awareness of the (moral) implications of a situation.	✳ I did not notice what was going on, or that something was amiss.	✳ I was aware of what was happening, or that something was amiss.	✳ I was aware of what was happening or that something was amiss. ✳ I had an understanding of the moral elements of the situation.
LOOK: understanding how emotions can help us to choose well.	✳ I was unable to identify my emotions, or the emotions of others.	✳ I was aware of my emotions and the emotions of others. ✳ I was not able to use awareness of these emotions to act in the right way.	✳ I was aware of my own emotions and those of others. ✳ I could use awareness of these emotions to try to act in the right way.

	STAGE 1	STAGE 2	STAGE 3
LISTEN: using reason to make deliberate (moral) choices.	✳ I could not identify the right thing to do in this situation. ✳ I did not know the middle way: the non-extreme actions. ✳ I could not avoid making the same mistakes I usually make in situations like this.	✳ I had an awareness of the right thing to do in this situation, but was unable to do it. ✳ I had an awareness of the middle way. ✳ I found it difficult to avoid making the same mistakes I usually make in situations like this.	✳ I knew the right thing to do in the situation. ✳ I could identify the middle way. ✳ I was aware of the mistakes I often make in situations like this and tried to avoid them.
CATERPILLAR: understanding of personal (moral) development.	✳ I am unaware of the impact of my actions on the person I might become. ✳ I would rather avoid pain or pursue pleasure than try to live life more skilfully.	✳ I was aware of the impact of my actions on the person I would like to become. ✳ My desire to avoid pain or pursue pleasure makes it very hard for to me to live skilfully.	✳ I was aware of the impact of my actions on the person I am trying to become. ✳ I can resist pain to live skilfully, but I find it harder to avoid pleasures.

As development in virtue becomes more sophisticated, we might see people progressing to stages 4, 5 and 6.

Figure 10: The caterpillar process: Stages 4-6

	STAGE 4	STAGE 5	STAGE 6
STOP	✳ I paused to assess the situation. ✳ I could avoid pain to do the right thing, but might have been swayed by pleasures; ending up doing the wrong thing, even though I knew what the right course of action was.	✳ I paused to carefully and deliberately assess the situation for its moral implications. ✳ I may have desired to avoid pain or pursue pleasure instead of the good, but I could fully overcome these desires.	✳ Stage 6 is like stage 5, but for the fully virtuous, doing the right thing happens without any effort of will: they take delight in doing the right thing.
NOTICE	✳ I was well aware of what was happening. ✳ I had a good understanding of the moral elements of the situation.	✳ I was fully aware of the situation. ✳ I had a full understanding of the moral elements of the situation.	
LOOK	✳ I understood my emotions and those of others. ✳ I could use those emotions to act in the right way.	✳ I fully understood my own emotional response and those of others. ✳ These emotions helped me to make the right choice about how to act.	

3.6

	STAGE 4	STAGE 5	STAGE 6
LISTEN	✳ I had a clear idea of the right thing to do in this situation. ✳ I had a clear idea of the middle way. ✳ I could avoid the mistakes I have made in situations like this in the past.	✳ I knew the right thing to do in this situation. ✳ I knew the middle way. ✳ I did not make the same mistakes I have made before, because I have learned from them.	
CATERPILLAR	✳ I am aware of the person I am trying to become and understood how my actions would affect that. ✳ I still find it hard to resist pleasures that hinder me from living really skilfully.	✳ I know what kind of person I am trying to become and I choose how to act carefully because of this. ✳ It is still sometimes an effort of will to do the right thing.	

3.6.4 Some other tools to reflect on student progress in character education

Reflective tasks

Perhaps the most simple and effective method of assessing the progress that students have made through a character education programme is to set them reflective tasks to complete (see section 4 for further information and suggestions of possible tasks). Character education materials typically involve increasing awareness of specific virtues coupled with practical strategies for ways of acquiring them. For example, there are certain performance virtues surrounding physical health, such as getting a good night's sleep, and there are associated processes which enable the acquisition of that virtue, such as switching off lights and screens in the run-up to bed time, or relaxation exercises. Reflective tasks will require students to analyse how effective these processes are for them and whether or not they have acquired those processes as habits. Reflective tasks are at their best when they focus on very specific elements of growth in virtue, such as the ability to pause before acting; the ability to read emotion; and the ability to identify the middle way between two extremes. These reflective tasks do not have to be written and inventiveness in ways of capturing such reflections is only limited by how far technology stretches. The materials produced can then be considered with reference to the character development grid (figures 9 and 10), with feedback given on ways of improving and developing. Consideration should be given to sub-section 3.3, which looks at the use of moral dilemmas; these can be a good way to identify the virtue reasoning skills of students.

Virtue spotting

Asking students to virtue spot on a regular basis can help to build their facility with naming the virtues and vices, and recognising them in action. This can be done in specific contexts such as friends/peers, family members, news stories, literature and TV/film, and can be very helpful in helping students to build not only a virtue/vice vocabulary, but developing the practice of seeing the virtues/vices in action and the impact that they can have.

Dialogue

In his book *Character Education* (2012), Wouter Sanderse describes how a process of Socratic dialogue can help students' character development. A Socratic dialogue involves two or more people sitting together and posing challenging questions in order to understand something more fully. By posing situations for students where the virtuous response is not obvious, and then challenging them and supporting them in refining their understanding, a Socratic dialogue can be used not only to assess progress, but to make progress too, especially in terms of deepening our own understanding in the light of viewpoints provided by others. This can be done across all four domains of virtue, from ways of solving technical or intellectual problems through to navigating moral issues where virtues such as loyalty and honesty come into conflict.

Student reports

Another method of tracking progress is to ask students to write their own reports about the progress they believe they have made in developing their character. Section 4 of this handbook has some additional suggestions as to how self-reflection by students can be structured. If students are able to build up a portfolio of reflective tasks, have had some experience of Socratic dialogue, and have navigated through some complex dilemmas, they may have some concrete evidence of how they have developed over a period of time. This should enable them to write a self-report paragraph to be published alongside other subject reports. This could be carried out in conjunction with reporting from form tutors, peers or family members.

3.6.5 Summary

This sub-section has outlined the boundaries that constrain character education evaluations. A framework to support students to learn and develop from self-reflection on their character development is also provided, along with some general activities to support the development of self-reflection. This framework can also be applied to section 4 of the handbook, which has examples of different approaches to self-reflection.

References

Ecclestone, K. and Hayes, D. (2009) *The Dangerous Rise of Therapeutic Education.* London, Routledge

Harrison, T., Bawden, M., and Rogerson, L (2016) *Teaching Character through Subjects*. University of Birmingham, Jubilee Centre for Character and Virtues [Online]. Available at http://www.jubileecentre.ac.uk/userfiles/jubileecentre/pdf/character-education/Teaching_Character_Through_Subjects.pdf (Accessed 17 August 2016)

Sanderse, W. (2012) *Character Education: A Neo-Aristotelian Approach to the Philosophy, Psychology and Education of Virtue*. Delft: Eburon Academic Publishers