Evaluation Checklist

The following evaluation checklist has been adapted from several existing sources, most notably work undertaken by the Centre for Social Mobility (2019) on behalf of the Office for Students (OfS), in addition to the reflections and expectations of Student Engagement, Evaluation and Research (STEER) at Sheffield Hallam.

Evaluation Criterion	Key Questions	Guiding Evidence
1. Initiative rationale	• What issue are you attempting to resolve and why? • Can you demonstrate an understanding of how and why your activity will lead to the desired outcomes? • Have you applied the SMART (Specific, Measurable, Achievable, Realistic and Time-specific) framework to your objectives?	Evaluators must demonstrate a strong understanding of: 1) what they are trying to achieve; 2) why an initiative is required to address the issue; 3) how the planned activity is expected to lead to the desired impact. This establishes a logical connection between the aim, outcomes and activities (Evaluation Support Scotland, 2019). **Developing a 'Theory of Change' can further show how an initiative contributes to a chain of results that enables its outcomes**, from short-term to longer-term, to be achieved (NPC, 2019). The **objectives should be heavily grounded in the rationale of the initiative**. The SMART framework is a simple and practical tool that helps evaluators set objectives that are clear, realistic and appropriate (Parsons, 2017).
2. Using evidence to inform initiative design	• What evidence do you have to support this approach? • How do you know that you have critiqued or gathered the best available evidence?	The **design of an initiative should be informed by existing evidence of 'what works'** to demonstrate the positive effect it is anticipated to have on participants. **In order to determine what evidence is most appropriate, evaluators need to reflect critically about the range of possible sources of evidence available** (Alliance for Useful Evidence, 2019), which should involve **challenging assumptions** and considering the contribution of 'less traditional' approaches (Austen & Jones-Devitt, 2019). An array of sources can be drawn upon to shape the design of an initiative, from the results of previous evaluations and relevant literature to anecdotal 'in the moment' reflections. In the case that there is limited or no evidence available, evaluation must be built in to generate evidence to inform how the activity is designed (Centre for Social Mobility, 2019).
3. Purpose of the evaluation	• Who are the key stakeholders of the evaluation? • What are you trying to find out from your evaluation? What is the hypothesis? • What aspects will be included within the focus of the evaluation and what aspects will be excluded?	**Two of the most common forms of evaluation are process and impact evaluations.** Process evaluations is concerned with the activity and outputs and focuses on 'evaluating the mechanisms through which an intervention takes place', such as how it has been implemented (Parsons, 2017, p. 16). Impact evaluations are concerned with measuring the consequential changes that results from those activities' (p. 17), such as the extent it is achieving its long-term goals. Many evaluations include both process and impact elements.

4. Indicators and measures	• Are you clear on how you will measure your outcomes at each stage of the evaluation (for example, short, medium and long-term)? • How will any unintended outcomes be detected and dealt with?	**For each outcome, it is necessary to select relevant indicators**, which are the measures that assess the changes that that the initiative is expected to make. Indicators should be specific, measurable, relevant to the planned activity (Centre for Social Mobility, 2019) and developed in consultation with stakeholders (Evaluation Support Scotland, 2019). It is recommended that indicators should be identified for each part of the evaluation process, with **consideration given to the role of short-term and medium-term outcomes as well as those that are longer-term**. The Kirkpatrick model is a useful way of identifying outcomes at four levels: Reaction; Learning; Behaviour; and Results (Centre for Social Mobility, 2019). It is necessary to consider how any **unintended consequences** or unexpected effect of activities, such as negative impact, will be taken into account (Parsons, 2017).
5. Evaluation design	• Which research method is most appropriate for the purpose of the evaluation? • Are you collecting data at different stages of the activity (before, after and during) to determine the magnitude of impact? • Are you using a comparison group to show what might have happened if the intervention had not taken place? • Are you utilising both quantitative and qualitative evidence?	In work commissioned for the OfS, the Centre for Social Mobility (2019) outlined three types of impact evaluation that generate different types of evidence: 1. A **narrative** standard demonstrates an understanding of knowing what you are doing and why, which may involve referring to existing evidence on the effectiveness of an initiative. 2. An **empirical enquiry** standard involves collecting data to measure changes generated by activities without establishing a causal effect. In order to reach the empirical standard, it is necessary to collect data: • to **demonstrate what would have happened if the initiative had not occurred**, which is known as the **counterfactual perspective**. The use of a comparison group, comprised of individuals with similar characteristics who did not take part in the activity, can help determine the effect of the initiative. • **at multiple points of the initiative** to provide a baseline measurement to help **determine whether there has been an improvement in outcomes over time**, such as between the start and the end of the activity and beyond. 3. The **causality** standard provides evidence of a causal effect of an initiative. A research design must be adopted that can establish this causal relationship on the outcome, such as an experimental or quasi-experimental design. The **design of the evaluation should be proportionate to what evaluators want to demonstrate, the existing evidence base and resources** (Centre for Social Mobility, 2019; Parsons, 2017). The narrative standard is the minimum requirement for all activities, while empirical or causality standards are expected for long-term or multi-activity interventions. **Adopting a range of quantitative and qualitative methods can be highly**

		complementary. Quantitative approaches can inform us 'about **what** happened, **where** and **when** and **who** with', while qualitative approaches 'helps us identify the factors or reasons affecting behaviour - the **how** and **why**' (Evaluation Support Scotland, 2019). Further guidance about the main methods and types of evidence used in Higher Education has been provided by Austen and Jones-Devitt (2019).
6. Data collection	• Who are you collecting data from? • What type of data is needed and when should you collect it? • Will the project be analysed at different levels (e.g. at an individual level, for specific groups of students, at a programme level)? • Can you make use of existing sources of evidence or do you need to collect new data? • If you have developed any new data collection tools, have they been piloted appropriately?	**Consider whether you need to collect new data or whether there is existing data that you can utilise.** If you are collecting new data, utilise pre-validated existing tools or ensure that any tools that you design on your own are piloted with participants, such as using cognitive interviewing. Examples of existing evidence are listed by Thomas (p. 32, 2017), such as data on key outcomes that institutions and bodies widely collect: • Survey data on the student experience, including the National Student Survey, Postgraduate Taught Experience Survey and the UK Engagement Survey. • Data provided by the Higher Education Statistics Agency (HESA) on UK performance indicators, such as widening participation, and on students and staff. **Benchmarking tools enable institutions to map their current practice** and to compare their data against the sector. **Contemplate whether there are other types of data collection besides a survey or a focus group that can meaningfully engage participants and generate informative data.** If you insist on using a survey, adhere to best practice to improve engagement and response rates. In the case of focus groups, consider ways that creative pedagogic approaches can be embedded.
7. Ethics	• Are procedures in place for addressing ethical considerations (e.g. to ensure no harm comes to participants) • If you are intending to publish your results externally, have you received ethical approval to do so? • How will data be stored in accordance with data protection guidance?	**All research and evaluation activity must adhere to guidance and best practice recommended by research councils and groups**, which relates to considerations for: participants, including avoidance of harm; stakeholders; and for dissemination (Parsons, 2017). At Sheffield Hallam, guidance on building ethical practice into institutional research and evaluation has been provided by the EthicsIR group at Sheffield Hallam (Austen, Chitwood & Jones-Devitt, 2019), while a new process of ethical approval is being implemented to allow institutional service evaluations to be shared externally. The **data of participants and stakeholders must be protected** by adhering to the requirements of General Data Protection Regulation (GDPR).
8. Planning the evaluation and resources	• Have you developed an evaluation action plan as part of the overall initiative specification? • What costs, time, knowledge,	**Evaluation should be structured and planned as part of the overall initiative specification**, rather than an ad-hoc activity (Parsons, 2017). **Completing a formal evaluation plan can enable an evaluation to be undertaken more effectively**. As part of the plan, Thomas (2017) suggests considering key questions such as:

	skills, or other resources are needed? Are these adequate? • Is any training required for capacity building purposes?	• What are the key activities, when will they take place and who will lead them? • What resources of support are required? • What will be the outputs of each activity? Existing frameworks are available for developing an evaluation, such as the RUFDATA (Reasons, Uses, Focus, Data, Audience, Timescale, Agency) tool. Ensuring that there is **sufficient resource and expertise to undertake the evaluation** is a fundamental consideration, as various roles and skills will be required. Establishing an institutional infrastructure and collaborating with staff who have experience of undertaking robust evaluations can help to build capacity.
9. Analysing and interpreting the results	• Are the conclusions appropriate for the quality of evidence that has been generated? • Have you triangulated findings from different sources to validate findings and to generate further understanding? • Have you drawn upon a range of perspectives?	The Centre for Social Mobility (2019) encourages evaluators to **consider whether the conclusions that they are making are appropriate for the quality of evidence that has been generated.** Guiding questions to consider include: • Is the conclusion reasonable and valid? • Does the study measure what it intended to measure? • Can the conclusion be generalised beyond its immediate context? **Develop further confidence in the conclusions by cross-verifying evidence from more than one source**, for example, combining quantitative data (e.g. survey) with quantitative information (e.g. interviews). **Draw upon data from a range of perspectives and consider using 'critical friends' and stakeholders to bring objectivity to the evaluation process**, for example, use of a steering group to review research findings and prioritise recommendations.
10. Use of evaluation	• Who is your audience(s) and what information do they need? • How will the results of the evaluations inform ongoing activities and future practice?	Evaluators have a responsibility to **consider how the learning from the evaluation will enhance practice** (Parsons, 2017), including to key stakeholders and those who contributed to the process. This requires identifying who the intended audience is and what mechanisms are in place to influence practice and enable a continuous improvement process to take place.
11. Sharing results	• Are there mechanisms in place to enable evaluation results to influence practice internally? • Are there mechanisms in place to influence practice externally across the sector?	Evaluators can have an active role in maximising the use of findings by **helping users to understand the evidence and its implications** (Parsons, 2017), for example, by supporting the dissemination of evidence. This might involve producing tailored research outputs to reach different audiences, for example: evaluation reports; briefing papers; presentations; and toolkits. **The findings of evaluations should always be made available, irrespective of whether the results are positive or negative**, so that they can inform the decision-making process. Evaluation findings must have received appropriate ethical approval for it to be used for external purposes.

Main References and Further Reading

Alliance for Useful Evidence (2019). *Using Research Evidence: A Practice Guide*. Retrieved from https://www.alliance4usefulevidence.org/assets/Using-Research-Evidence-for-Success-A-Practice-Guide-v6-web.pdf (last accessed 18th November 2019)

Austen, L., Chitwood, E., & Jones-Devitt, S. (2019). *Building Ethical Practice Into Institutional Research and Evaluation*. Retrieve from https://blogs.shu.ac.uk/steer/files/2019/10/ETHICS-IR-BOOKLET-Draft-Five.pdf (last accessed 18th November 2019)

Austen, L.& Jones-Devitt, S. (2019). *Guide to Using Evidence*. Retrieved from https://www.enhancementthemes.ac.uk/docs/ethemes/evidence-for-enhancement/guide-to-using-evidence.pdf?sfvrsn=394c981_14 (last accessed 18th November 2019)

British Educational Research Association [BERA] (2018). *Ethical Guidelines for Educational Research*, fourth edition, London. Retrieved from https://www.bera.ac.uk/researchers-resources/publications/ethical-guidelines-for-educational-research-2018 (last accessed 18th November 2019)

Centre for Social Mobility (2019). *Using standards of evidence to evaluate impact of outreach*. Retrieved from https://www.officeforstudents.org.uk/media/f2424bc6-38d5-446c-881e-f4f54b73c2bc/using-standards-of-evidence-to-evaluate-impact-of-outreach.pdf (last accessed 18th November 2019)

Evaluation Support Scotland (2019). *ESS Support Guides*. Retrieved from: http://www.evaluationsupportscotland.org.uk/resources/ (last accessed 18th November 2019)

Higher Education Statistics Agency [HESA] (2019). *Open data and official statistics*. Retrieved from https://www.hesa.ac.uk/data-and-analysis (last accessed 18th November 2019)

Leadership Foundation for Higher Education (2017). *Research Leaders Impact Toolkit: Evaluating*. Retrieved from https://www.lfhe.ac.uk/en/research-resources/publications-hub/index.cfm/RLITEvaluating (last accessed 18th November 2019)

National Foundation for Educational Research (2019). *Research Methods & Operations*. Retrieved from https://www.nfer.ac.uk/key-topics-expertise/research-methods-operations/ (last accessed 18th November 2019)

NPC (2019). *Theory of change in ten steps*. Retrieved from: https://www.thinknpc.org/resource-hub/ten-steps/ (last accessed 18th November 2019)

Nutley, S., Powell, A., & Davies, H. (2013). *Provocation Paper for the Alliance for Useful Evidence*. Retrieved from: https://www.alliance4usefulevidence.org/assets/What-Counts-as-Good-Evidence-WEB.pdf (last accessed 18th November 2019)

Office for Students [OfS] (2019). *A to Z of effective practice in access and participation*. Retrieved from https://www.officeforstudents.org.uk/advice-and-guidance/promoting-equal-opportunities/evaluation-and-effective-practice/a-to-z-of-effective-practice-in-access-and-participation/ (last accessed 18th November 2019)

Parsons, D. (2017). *Demystifying evaluation: Practical approaches for researchers and users*. Policy Press.

Thomas, L. (2017). Evaluating student engagement activity: Report, evaluation framework and guidance. The Student Engagement Partnership. Retrieved from: http://tsep.org.uk/wp-content/uploads/2017/06/Student-Engagement-Evaluation-Framework-and-Report.pdf (last accessed 18th November 2019).

Student Engagement, Evaluation and Research [STEER]. *Evaluation*. Retrieved from https://blogs.shu.ac.uk/steer/evaluation/ (last accessed 18th November 2019)

Produced by Directorate of Student Engagement, Evaluation and Research, Sheffield Hallam University 2019