Tools and Methods for Evaluation

Career Internship Network Publications

The Career Internship Network (CIN), a program of the Youth Development Institute, comprises more than thirty New York City institutions with established internship programs for adolescents that provide exposure to, understanding of, and experiences in potential careers. As part of its services, CIN provides professional development workshops for members who run the career development programs. Resource guides and worksheets accompany each workshop to assist program coordinators in their work with interns. This downloadable publication is part of a series adapted from these workshop materials that we want to share with the field. We encourage you to use any of the materials, but please include copyright information on related pages, if any appear.

Tools and Methods for Evaluation
Sound evaluation takes more than simply conducting surveys or questionnaires and eye-balling the information. Before you gather your data you need to put in a good deal of careful thought and planning to decide what tool you will use for getting data and/or the method of gathering it. *Tools* are *tangible* objects such as journals or questionnaires. *Methods* are *processes* for gathering data such as holding a focus group or interviewing participants. There are benefits and challenges to each tool or method and you need to take your resources, time, and personnel into account when deciding which to use.

When developing tools and/or planning methods for evaluation you need to start out by determining what *questions* you have that you want answered and the *data* that will provide the necessary information. From there you can set up *how* you will gain your data. You also need to consider the *kind of data* each tool or method will gather. It is very important to pre-test your tools and/or methods to make sure they are getting the type of information you need. Finally, sound evaluation should involve the combination of at least *three* tools and/or methods and preferably *three* people to analyze and interpret the information. The raters can be either/or/combination of insiders or outsiders of your program. One of the benefits of insiders is cost effectiveness. However, outside evaluators can be more objective about the data collected and its meaning. A combination of stakeholders and outside consultants provide benefits—the consultant can keep the evaluation on track and the stakeholders can provide multiple perspectives on the evaluation process.

Evaluation and Assessment
Methods

Methods	Description	Benefit	Drawbacks
Anecdotal Information	Can be gathered through observation, interviews, study of videotapes and other media sources. It can be reported "raw" or after organizing the data to reveal behavioral, cognitive or social patterns and relationships. It is generally a time consuming multi-step process that is useful in providing the context for teaching and learning	• In moment • Candid • Removes anxiety (both sides) • Takes learning styles into account • Easy and inexpensive • Raw	• Non-quantifiable • Random • Biased • Situational • Hard to code • Accuracy of accounts
Authentic Reflection Stems/Sentence Starters	Method, typically writing, that uses three parts of a sentence starter to get deeper responses from participants. (e.g., *The best thing about the internship was...because...What I will come away with from the program is...*) The intention of authentic reflection is to generate a deeper, more thoughtful and nuanced response	• Deepens understanding of responder to experience • More nuanced information • Can be used to make formative changes to program • Process of responding can help responder, practitioner and evaluator alike • Exact words from the responder (first hand voice) can be used with great impact in reports and proposals	• Must test authentic reflection stems to make sure that they are not skewing responder to particular answer • Time consuming to analyze • Should incorporate more than one person (ideally three or more) in analyzing and interpreting results • Time consuming to administer
Case Studies	The collection and presentation of detailed information about a particular participant or small group, frequently including the accounts of subjects themselves.	• Up close and personal perspective of person's experience • Gathered over time	• Time consuming • Can make mistaken generalizations from fairly small number of people or small group being studied • Interpretation can be biased
Focus Groups	Groups that are brought together randomly to provide their opinions and attitudes on a set of specific questions	• Room to gather a lot of data (through discussion) • Broad range of participants input • Participants gain a sense of value	• Time consuming • Dependent on participants attendance • Need a facilitator and a recorder to gather data (training required)

Methods	Description	Benefit	Drawbacks
Interviews (Pre or Post)	Structured (or free form) conversations that can capture detailed information about the experience of a project, and may encourage expression of opinions, feelings, ideas, and understandings	• Ability to see the before and after results • Easy to administer • Open-ended • More expressive verbally • Emphasizes importance of students feedback	• Difficult to analyze data (there's room for misinterpretation) • Hard to be honest if (-) feedback – candor • Nervous! • How much can be captured • (rely on memory? Notes?) • Consistency of topics covered
Observation	Method that includes systematic, purposeful study of what is happening to an individual or group. Notes are jotted and transcribed (in a variety of ways) either on the scene or immediately after, from memory; and/or an observation checklist (see below) is used	• Captures a reality • Real-time • Promotes discussion • Captures nuance outside eval forms • On going observation vs. • A lot of useful data is gathered • Easy to administer • Opportunity to learn about the teens different learning styles, see them in action	• Very subjective • Can be variable (off-day..) • Consistency (observing each student under similar conditions) • Per formative • "observation day" • Data gathered can be biased • Difficult to analyze and interpret data
Performance Task Assessment	Determination of the quality of someone doing/completing an activity that demonstrates participant's ability to apply acquired skills on tasks for which the skills are appropriate (e.g., a performance task might be the ability to build a boat, run a workshop that ultimately creates a tour, etc.)	• Concrete, demonstrates comprehension • Opportunity for participant to improve • Gives a chance to put skills into action • Can build on the team component	• Time: observe each student • Multiple perspectives needed to evaluate • Outside distractions • Time-sensitive (a facilitator needed at the time of performance)
Portfolio Assessment	A portfolio is collection of work, usually drawn from students' classroom work. A portfolio becomes a portfolio assessment when (1) the assessment purpose is defined; (2) criteria or methods are made	• Self-learning tool • Creative, engaging • Gives a sense of accomplishment to see final portfolio • Provides future benefits	• Needs to be structured • Requires a facilitator to guide teens by providing input • Can be overwhelming to put the portfolio together

Methods	Description	Benefit	Drawbacks
Portfolio Assessment (cont.)	clear for determining *what* is put into the portfolio, *by whom*, and *when*; and (3) criteria for assessing either the collection or individual pieces of work are identified and used to make judgments about performance. Portfolios can be designed to assess student progress, effort, and/or achievement, and encourage students to reflect on their learning.	- Concrete body of work - Self-expressive - Appeals to varied styles	- Time consuming (to create, assess) - Space-consuming
Random Sampling	A process that uses a sample of a population selecting it strictly by chance, yielding no discernible pattern beyond chance.[1]	- Can evaluate a larger group with smaller sample - Makes manageable - Save resources	- Time consuming - Process of doing genuine sampling selection can be challenging to novices - Can result in skewed results

---

[1] Random sampling can be accomplished by first numbering the population, then selecting the sample according to a table of random numbers or using a random-number computer generator. The sample is said to be random because there is no regular or discernible pattern or order. A less rigorous method of random sampling is putting everyone's name in a hat and randomingly selecting the number you want for your sample group. Random sample selection is used under the assumption that sufficiently large samples assigned randomly will exhibit a distribution comparable to that of the population from which the sample is drawn.

Evaluation and Assessment
Tools

Tool	Description	Benefits	Drawbacks
Examinations/Tests	A set of questions or problems designed to assess knowledge, skills, or intelligence	• Quantitative • Specific answers	• Only gets information about what you ask • Limited information about impact on the student • Failure vs. success • Biased
Likert Scales	Evaluation/assessment tool to gather opinions or attitudes about specific questions related to a particular belief, experience, or event. Generally structured across a continuum (scale) with numbers of say, 1 – 4, with phrases defining their value, e.g., from "don't agree at all" to "completely agree."	• Quantitative • Fast • Easy to interpret	• Doesn't provide reason why a score was selected/people might have different reasons for picking a similar number thus creating the need for an articulation of what each number means: a rubric • Rubrics can be time consuming to create • Rubrics need to be tested across observers and observees
Learning Logs and Journals	Strategies for collecting data by the researcher or the participants and when formally studied and analyzed, can reveal valuable patterns of data or information about what is being learned. Learning logs are short and descriptive, whereas reflective journals are more descriptive and have to do more with reflection and connection, response to literature	• Gets useful information that you might not have been looking for • More nuanced • Open ended, own words • Self-learning tool	• More time consuming to administer, read, analyze and interpret (and respond to if appropriate)
Observation Checklist	Strategy to assess specific skills, behaviors, or dispositions (of single participant or of an entire group)	• Clear information • Unbiased • Provides a focus • Immediate feedback • Organized • Creates baseline	• Can be stressful to observe • Time consuming • Subjective • Biased • Because goals are predetermined, it can not be as responsive to

Tool	Description	Benefits	Drawbacks
Observation Checklist (cont.)		• Good in the moment • Validates importance of human moment and interpretation • Observer can learn and reflect as tool is applied	• anomalies • Behavior of group can change when knowing it is being observed • Tools must be used with pre and post reflection to provide any depth of interpretation
Questionnaires	Sheets with questions that can capture binary (yes or no) answers to questions that can be counted, tallied and interpreted	• Can get the information you want • Consistent • Precise data • Easy • Fast • Good for simple assessment yes/no	• Too vague or broad • Confusing questions • Question set up can bias responders to particular answer • Miss vital (other) information • Answers may not always reflect knowledge, beliefs, or feelings • Responders might complete them fast without much thought • Can be limiting • Specific, but not authentic • For good evaluation or assessment, need other tools as well (detailed)
Rubrics	Tool to help define *quality* or *achievement*, or *performance.* Rubrics (1) Identify the key *criteria*/*components* of what is being assessed/evaluated and; (2) They likewise *describe gradations* of quality for *each* criterion/component from "not acceptable" ("not observed," etc.) " to "outstanding."	• Gets everyone on the same page • This is what we want to go for	• You really have to know what you want to measure • Time consuming • Detailed process
Surveys (Written and Verbal)	A tool in which responders give answers about the condition, situation, or value of something	• Can get at specific info • With technology, can reach a greater audience and record data • Time efficient • Can be qualitative • Anonymous can be a good thing • Online	• Will become "yes" or "no" • Respondents can be harsh • May not answer the question you are asking •