**Student Affairs Assessment Improvement Rubric**

Contributors in alphabetical order:

Combs, Aaron

Finney, Sara

LeRoy, Stephanie

McCleve, Michael

Newman, Misty

Purks, Tripp

Stewart, Jonathan

Wild, Autumn

Wilson, Saleena

Center for Assessment and Research Studies &

Student Affairs Assessment Advisory Council

James Madison University

Latest Version: February 2024

The Student Affairs Assessment Improvement Rubric © 2024 by Finney, S., Stewart, J. & Wild, A. is licensed under [Attribution-NonCommercial-ShareAlike 4.0 International](https://nam10.safelinks.protection.outlook.com/?url=http%3A%2F%2Fcreativecommons.org%2Flicenses%2Fby-nc-sa%2F4.0%2F%3Fref%3Dchooser-v1&data=05%7C02%7Cfinneysj%40jmu.edu%7C666d6cbd3c474238b1f108dc21a4ec16%7Ce9333c23cac742f499895cee3d4a79c0%7C0%7C0%7C638422239412587907%7CUnknown%7CTWFpbGZsb3d8eyJWIjoiMC4wLjAwMDAiLCJQIjoiV2luMzIiLCJBTiI6Ik1haWwiLCJXVCI6Mn0%3D%7C0%7C%7C%7C&sdata=I0wiZShP65wjGc7zXQjx8d0Efj0hvcvXpWuc93Ettw0%3D&reserved=0)

**Criterion 1 – Student Learning & Development Outcomes**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Does the report include student learning outcomes? | | | | | | | | |
| **If yes, then …** | | Complete current block | | **If no, then …** | | | Provide an “average score” of zero, and move to next criterion | |
|  | | | | | | | | |
| **Desired Student Learning & Development Outcomes** | Precise statements of what students will know, think, or be able to do because of programming. | | | | | | |
| **Sub-Criteria** | **Exemplary (3)** | | **Proficient (2)** | | **Developing (1)** | **Missing (0)** | | **Score** |
| Audience | **All** outcomes are stated in student-centered terms. | | **Most** outcomes are stated in student-centered terms. | | **Some** outcomes are stated in student-centered terms. | **No** outcomes are stated in student-centered terms. | |  |
| Measurable | Measurable using **precise verbs** (e.g., apply, describe, commit to, accept, indicate, display, agree to, identify) that **detail the desired degree** of observable knowledge, attitudes, and/or behaviors (e.g., increase 1/2 standard deviation from before to after program completion; will score about 80% correct; will achieve a 2.0 GPA). | | Measurable using **precise verbs** (e.g., apply, describe, commit to, accept, indicate, display, agree to, identify) that detail observable knowledge, attitudes, and/or behaviors, but **no noted desirable degree** (e.g., increase from before to after program completion) is mentioned. | | Not measurable due to **use of imprecise verbs** (e.g., know, understand). | Not measurable due to **lack of verbs** that indicate observable knowledge, attitudes, and/or behaviors. | |  |
| Malleable | **All** outcomes represent malleable knowledge, attitudes, and/or behaviors that are sensitive to programming and **references are included** to support these claims. | | **Most** outcomes have research supporting their malleability. | | **Few** outcomes have research supporting their malleability. | It is **unclear** if the outcomes are malleable when employing this programming in the given context. Or **no** mention of malleability of outcomes. | |  |
| Equity Centered | Outcomes are **relevant** for targeted student populations and there is **evidence to** **support** those claims. **Students were** **included** in the creation or revision of outcomes. | | Outcomes are **relevant** for targeted student populations and there is **evidence to** **support** those claims. | | **Unclear if outcomes are relevant** for targeted student populations, and this lack of evidence/knowledge is **addressed as a limitation**. | **No mention** of equity issues related to outcomes. | |  |
| Comments: |  | | | | | Average Score: | |  |

|  |  |
| --- | --- |
| Sub-Criterion Definitions: |  |
| Audience | The audience refers to the group that the facilitators/educators intend to impact through the planned programming. For example, if the program is relevant for first-year students at the university, the audience is the first-year students who participated in the program. |
| Measurable | Measurable, in this instance, indicates that an observer (e.g., student affairs educator) can gauge levels of the outcome using empirical evidence. Verbs that support measurement of the outcome are often mentioned in this context. These verbs specify an observable action on the part of the learner. Measurable verbs help to describe and classify observable knowledge, skills, attitudes, values and behavior. |
| Malleable | A malleable outcome represents knowledge (e.g., intercultural competence), attitudes (e.g., sense of belonging), and/or behaviors (e.g., civic engagement) that can be changed with an intervention/program/experience. The outcome is not trait-like (e.g., agreeableness). |

**Criterion 2 – Program Theory**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Does the report include program theory? | | | | | | | | |
| **If yes, then …** | | Complete current block | | **If no, then …** | | Provide an “average score” of zero, and move to next criterion | | |
|  | | | | | | | | |
| **Using Theory and Research to Create and/or Map Programming to the Outcomes** | A synopsis of the specific theory and research used in the creation and mapping of each programming element is provided. | | | | | | |
| **Sub-Criteria** | **Exemplary (3)** | | **Proficient (2)** | | **Developing (1)** | | **Missing (0)** | **Score** |
| Outcome-Programming Mapping | **All** program elements are mapped to their respective outcomes. | | **Most** program elements are mapped to their respective outcomes. | | **Some** program elements are mapped to their respective outcomes. | | **No** program elements are mapped to respective outcomes or elements are **not** clearly identified. |  |
| Program Theory | A theoretical framework is identified, cited, and applied to **all** elements of the program (mapping between programming and outcomes is **clear** and **justified**). | | A theoretical framework is identified, cited, and applied to **most** of the elements of the program. | | A theoretical framework is identified, cited, and applied to **some** of the elements of the program. | | A theoretical framework is **not** identified, cited, and applied to any element of the program. |  |
| Evidence-Based Practice | Evidence-based practices have been articulated and provide rationale for **all** program elements. | | Evidence-based practices have been articulated and provide rationale for **most** of program elements. | | Evidence-based practices have been articulated and provide rationale for **some** of the program elements. | | Evidence-based practices have **not** been articulated and **do** **not** provide rationale for any program elements. |  |
| Reasonable | Outcomes are **reasonable** given length and strength of the program and reasonableness is **extensively** **supported** **by** **evidence** from previous research. | | Outcomes are **reasonable** given length and strength of the program and reasonableness is **minimally** **supported** **by** **evidence** from previous research. | | It is **unclear** if the outcomes are reasonable given the information provided in the report. | | Outcomes are **not** reasonable given the length and strength of the program. |  |
| Equity Centered | **Clear** articulation about if and how the etiology of outcomes differs across student populations.  Intersectionality of students’ identities is **considered** and **addressed**. **Student feedback is solicited** on the clarity of the applied program theory. | | **Clear** articulation about if and how etiology of outcomes differs across broad student populations. | | **Unclear** or **limited** articulation around the differences in etiology across student populations. **Acknowledgement** that this assessment process may be the first step in identifying these differences. | | **No** discussion of potential differences in etiology |  |
| Comments: |  | | | | | | Average Score: |  |

|  |  |
| --- | --- |
| Sub-Criterion Definitions: |  |
| Outcome-Programming Mapping | Outcome-program mapping is a diagram or table showing how each outcome is addressed by program elements (e.g., activities, strategies, readings). |
| Program Theory | Program theory explains how and why programming is expected to be effective. Why should engaging in these programming elements (e.g., activities, content, pedagogy) result in the stated outcomes for this population of students? |
| Evidence-Based Practice | Evidence-based practice refers to programming that has been empirically evaluated and been shown to effectively impact intended outcomes. |
| Reasonable | Reasonable outcomes are feasible for the program and population being served. Some outcomes cannot be achieved given university resources. |

**Criterion 3 – Outcome Measures**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Does the report include information about the measures of the outcomes? | | | | | | | |
| **If yes, then …** | | Complete current block | | **If no, then …** | | Provide an “average score” of zero, and move to next criterion | |
|  | | | | | | | |
| **Screening Question:** | * If measures were all selected from existing measures, rate “Measure Selection,” “Use of Direct Measures” & “Measures-to-Outcomes Map.” Skip “Measure Development” and provide no rating. * If measures were all developed, rate “Measure Development,” “Use of Direct Measures” & “Measures-to-Outcomes Map.” Skip “Measure Selection” and provide no rating. * If measures were both selected and developed, rate all sub-criteria below. | | | | | | |
|  | | | | | | |  |
| **Selecting or Creating Measures of the Outcome** | A description of how outcomes are operationalized is provided. Often discussed as "tools", "measures" or "instruments" and include self-report measures, tests, rubrics (to rate writing, performances, portfolios, products, presentations, etc.), or observation tools (e.g., did a student engage in particular behaviors?). A direct link should be made between specific measures and stated outcomes. | | | | | |  |
| **Sub-Criteria** | **Exemplary (3)** | | **Proficient (2)** | | **Developing (1)** | **Missing (0)** | **Score** |
| Measure Selection | Pre-existing measures with **high-quality** psychometric properties have been chosen from the literature and those properties are **reported**. | | Pre-existing measures with **adequate** psychometric properties have been chosen from the literature and those properties are **reported**. | | Pre-existing measures with **poor** psychometric properties have been chosen from the literature and those properties are **reported**. | Pre-existing measures were chosen but **no** psychometric properties presented. |  |
| Measure Development | Rigorous measure development process is detailed, and validity and reliability **evidence is reported.** | | Rigorous measure development process is detailed, but **no** **evidence** of psychometric properties. | | **Unclear** how the measure was developed. |  |  |
| Use of Direct Measures | **All** outcomes are assessed with direct measures of the outcome (measures reflect what students know, value, and can do). | | **Most** outcomes are assessed with direct measures of the outcome. | | **Some** outcomes are assessed with direct measures of the outcome. | **No** outcomes are assessed with direct measures of the outcome (e.g., satisfaction or attendance used to infer learning or development). |  |
| Measures-to-Outcomes Map | **All** outcomes are mapped to measures that represent the outcome and the match between the two is explained in detail (e.g., reviewed several existing measures and the selected measure aligns best with outcome; designed a measure and had others engage in backwards translation to confirm match between measure & outcome). | | **Most** outcomes are mapped to measures with their match being explained in detail; for other outcomes, the details are brief or vague regarding the match. | | **Some** outcomes are mapped to measures with their match being explained in detail; for other outcomes, the details are brief or vague regarding the match. | **Superficial** match between measures and outcomes. |  |
| Equity Centered | The measure, whether developed or selected, **produces** scores that allow for equally trustworthy inferences about the outcome, regardless of student population. **Multiple forms of evidence** are provided to support this claim. **Student feedback is** **solicited** regarding clarity of the measures used. | | The measure, whether developed or selected, ***should*** **produce** scores that allow for equally trustworthy inferences about the outcome, regardless of student population. **Limited** **evidence** is available or provided to support this claim. | | There is **no evidence** that the measure functions equivalently across different student populations and because of this, there is **language cautioning** the use of the scores to make inferences. | **No mention** of how the measure functions across different student populations. |  |
| Comments: |  | | | | | Average Score: |  |

**Criterion 4 – Implementation Fidelity**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Does the report include implementation fidelity information? | | | | | | | |
| **If yes, then …** | | Complete current block | | **If no, then …** | | Provide an “average score” of zero, and move to next criterion | |
|  | | | | | | |  |
| **Implementation Fidelity Data** | A description of the alignment between the planned programming (e.g., curriculum, pedagogy, activities, strategies) and the implemented programming (i.e., the programming the students experienced). | | | | | |  |
| **Sub-Criteria** | **Exemplary (3)** | | **Proficient (2)** | | **Developing (1)** | **Missing (0)** | **Score** |
| Inclusion of Core Aspects of Implementation Fidelity | The implementation fidelity checklist **includes** **and** **specifies** ***all*** core aspects: program differentiation, adherence, exposure, quality, and responsiveness. | | The implementation fidelity checklist **includes but does *not* clearly specify** all core aspects: program differentiation, adherence, exposure, quality, and responsiveness. | | The implementation fidelity checklist **includes some, but not all,** core aspects. | The implementation fidelity checklist **does *not* include any** core aspects of implementation fidelity. |  |
| Clarity of Description of Implementation Fidelity Methods | Provides **clear** **descriptions** of methods used to collect implementation fidelity (from who, with what, and when). Someone can **replicate** this data collection. | | Provides **descriptions** of the methods used to collect implementation fidelity with some indication of who, what, or when. | | Provides a **vague/unclear** **description** of the methods used to collect implementation fidelity data. **Cannot be replicated.** | Provides **no descriptions** of the methods used to collect implementation fidelity data. |  |
| Description of  Implementation Fidelity Results | For **all** outcomes, clearly specifies if all programming was implemented as planned, most programming was implemented as planned (noting which programming was not implemented as planned), some programming was implemented as planned (noting which programming), or no programming was implemented as planned. | | For **some** outcomes, clearly specifies if all programming was implemented as planned, most programming was implemented as planned (noting which programming was not implemented as planned), some programming was implemented as planned (noting which programming), or no programming was implemented as planned. | | **General** **statements** across outcomes regarding the implementation of programming. | **No discussion** of the implementation fidelity results. |  |
| Use of  Implementation Fidelity Data | Implementation fidelity data **accurately** informs the evaluation of the programming’s effectiveness **and recommendations** for its improvement. | | Implementation fidelity data **accurately** informs the evaluation of the programming’s effectiveness. | | There was an attempt to use implementation fidelity data to inform the evaluation of the programming’s effectiveness, but it was **not accurately and/or clearly** interpreted. | Implementation fidelity data was **not** used to inform the evaluation of the programming’s effectiveness. |  |
| Equity Centered | Examined whether program components were implemented with equal fidelity across diverse student populations and findings are **clearly** **articulated** in the report.  Plans to address implementation discrepancies are **included, if applicable**.  Students are **involved** in gathering or interpreting implementation fidelity data. | | Examined whether the program components were implemented with equal fidelity across diverse student populations and findings are **clearly** **articulated** in the report. | | **Crude** data collection is used to make claims about whether the program components were implemented with equal fidelity across diverse student populations. | **No mention** of whether the program components were implemented with equal fidelity across diverse student populations. |  |
| Comments: |  | | | | | Average Score: |  |

**Criterion 5 – Collecting Outcomes Data**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Does the report include information about the collection of outcomes data? | | | | | | | | |
| **If yes, then …** | | Complete current block | | **If no, then …** | | Provide an “average score” of zero, and move to next criterion | | |
|  | | | | | | | |  |
| **Collecting Outcomes Data** | A description providing clear information on the process of gathering outcomes data is included. This description allows others to judge the veracity of the process and to replicate data collection if necessary. | | | | | | |  |
| **Sub-Criteria** | **Exemplary (3)** | | **Proficient (2)** | | **Developing (1)** | | **Missing (0)** | **Score** |
| Data Collection Design | Data collection design was appropriate for **all** intended inferences, as articulated in the SLOs. | | Data collection design was appropriate for **most** of the intended inferences, as articulated in the SLOs. | | Data collection design was appropriate for **some** of the intended inferences, as articulated in the SLOs. | | Data collection design was **not** **appropriate** for any of the intended inferences, as articulated in the SLOs. |  |
| Replicability | Data collection process was detailed so **precise replication is feasible**. Details include: description of sample, data collection protocol, data collection context, student motivation in data collection. | | Data collection process was detailed to the extent that **approximate replication is feasible.** | | **Basic** information about data collection is provided (who and how many students engaged in providing outcomes data), but this information is **not enough** to judge the veracity of the process or replicate the process. | | Data collection details **not provided.** |  |
| Threats to  Validity of Inferences | Threats to validity are **explained** and **proactively protected against**. | | Threats to validity are **explained** and **mostly protected against.** | | Threats to validity are **explained** but **not protected against**. | | Threats to validity are **not** **mentioned**. |  |
| Equity Centered | **Clear** articulation of how the following impacted the equity of the outcomes data collection process:  1) power & positionality of assessor  2) data collection method (multiple methods used, Universal Design for Learning considered, etc.)  3) institutional climate and culture (e.g., what are students’ perceptions of engaging in data collection?)  Plans to address negative impact of these effects are provided **in detail**.  **Students involved** in data collection. | | A **clear** articulation of how the following impacted the equity of the data collection process:  1) power & positionality of assessor  2) data collection method  3) institutional climate and culture | | **Unclear or limited** articulation of how the following impacted the equity of the data collection process:  1) power & positionality of assessor  2) data collection method  3) institutional climate and culture | | There is **no mention** of how the following impacted the equity of the data collection process:  1) power & positionality of assessor  2) data collection method  3) institutional climate and culture |  |
| Comments: |  | | | | | | Average Score: |  |

|  |  |
| --- | --- |
| Sub-Criterion Definitions: |  |
| Data Collection Design | A data collection plan supports the inferences educators intend to make (as stated in the outcomes). If outcomes specify “change” or “growth,” a longitudinal design which assesses outcomes before and after programming is necessary. If outcomes specify that students experiencing the programming will know, value, or perform better than students not experiencing the programming, a design that includes a comparison group is necessary. |
| Replicability | The quality of being able to repeat the data collection process and reproduce the same/similar results. |
| Threats to Validity | Threats to validity are factors that impact the inferences educators can make from assessment results. Threats to validity impact the extent to which educators can attribute findings to the program (i.e., program caused students to increase in desired outcome). Threats to validity include external factors that may have affected results (e.g., attrition of students from program or data collection; natural maturation of students; history effects such as a pandemic, suicide on campus, election results). |

**Criterion 6 – Data Analysis**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Does the report include data analysis? | | | | | | | | |
| **If yes, then …** | | Complete current block | | **If no, then …** | | Provide an “average score” of zero, and move to next criterion | | |
|  | | | | | | | |  |
| **Analysis of Data** | A complete and appropriate analysis of the data is accurately interpreted for each outcome. | | | | | | |  |
| **Sub-Criteria** | **Exemplary (3)** | | **Proficient (2)** | | **Developing (1)** | | **Missing (0)** | **Score** |
| Appropriateness of Analyses | **All** outcomes are evaluated using appropriate analyses. | | **Most** outcomes are evaluated using appropriate analyses. | | **Some** outcomes are evaluated using appropriate analyses. | | **No** outcomes are evaluated using appropriate analyses. |  |
| Interpretation of Analyses | Interpretations of all presented results are **provided**, and the interpretations are **accurate**.  Interpretations conducted by **at least 2 people.** | | Interpretations of all presented results are **provided**, and interpretations are **accurate**. | | Interpretations were **attempted** but were **not accurate** (not supported by results) or **not complete** (some results are presented but not interpreted). | | **No interpretation** of analyses was attempted. |  |
| Equity Centered | Data are **intentionally disaggregated** to reflect existing student populations and analyzed by **multiple individuals**, such as faculty and students, to examine differential program effectiveness across student populations.  **Care is taken to avoid language or assumptions** that every student in a sub-population is the same (focus on variability as much as averages).  A method to represent findings from small samples is **employed** and **explained** (i.e., small populations are *not* ignored, but analyses differ from larger populations).  **Students are invited to engage in the data analysis process**. | | Data are **intentionally disaggregated** to reflect the existing student populations and analyzed (**unclear** by how many individuals) to examine differential program effectiveness across student populations.  **Care is taken to avoid language or assumptions** that every student in a sub-population is the same. | | Data are disaggregated **crudely** and **analyzed** to examine differential program effectiveness across broad student populations. | | There is **no disaggregation** of data by student populations. |  |
| Comments: |  | | | | | | Average Score: |  |

|  |  |
| --- | --- |
| Sub-Criterion Definitions: |  |
| Appropriateness  (In terms of analyses) | An appropriate analysis aligns with the question being asked and the data collected. For example, if the program aims to increase effective student leadership over a month-long camp and utilizes a pre-post design, it would be appropriate to use a repeated measures t-test. |
| Interpretation  (In terms of analyses) | The interpretation of an analysis refers to the inferences/conclusions that can be made as a result of the analysis used. |
| Data Disaggregation | Conducting analyses and reporting results for different student populations. Disaggregation is necessary, but one must consider the level of disaggregation. Crude disaggregation can lead to a false representation of “homogenous” groups (e.g., disaggregating students by broad ethnic categories, broad disability categories). Must provide demographic variable options that reflect the student populations to disaggregate appropriately. |

**Criterion 7 – Reporting Results**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Does the report include results? | | | | | | | | |
| **If yes, then …** | | Complete current block | | **If no, then …** | | Provide an “average score” of zero, and move to next criterion | | |
|  | | | | | | | |  |
| **Reporting Results** | A coherent summary of the results is situated in previous findings and useful for a variety of stakeholders. | | | | | | |  |
| **Sub-Criteria** | **Exemplary (3)** | | **Proficient (2)** | | **Developing (1)** | | **Missing (0)** | **Score** |
| Achievement of Desired Outcomes | **Notes** and **explains** whether each outcome was met, partially met, or not met. | | Only **notes** (**does not explain**) whether each outcome was met, partially met, or not met. | | **Selectively reports** whether outcomes were met, partially met, or not met. | | **Does not address** whether any outcomes were met, partially met, or not met. |  |
| Comparison of Findings to Previous Reports  *(Skip if there is no*  *previous report)* | Past trends/previous results are **provided**, **explained**, and **compared** to current results. | | Past trends/previous results are **provided** and **explained** but are **not compared** to current results. | | Past trends/previous results are **provided** but **not explained** and **not compared** to current results. | | Past trends/previous results are **not provided**. |  |
| Sharing Results | Results are shared with stakeholders (e.g., educators, facilitators, students, upper administration), stakeholders are **clearly** **identified**, and method of communication is **explained** (e.g., division meeting, shared drive, website). | | Results are shared with **identified** stakeholders, but the method of communication is **not provided/explained** | | Results are **shared**, but it is **unclear with whom and how**. | | Results are **not shared (no mention of sharing).** |  |
| Equity Centered | Presentation of results **ensures** representation of relevant student populations.  Deficit-based wording is **avoided** when reporting results.  Relevant student populations are **invited to contribute** to the writing of the report.  The dominant/majority group is **not used** as the norm for comparison. | | Presentation of the results **ensures** the representation of relevant student populations.  Deficit-based wording is **avoided** when reporting results. | | There is an acknowledgement within the presentation of results that there **may be differential effectiveness** across student populations, but there is **no formal exploration** of equity issues regarding results. | | There is **no discussion** of equity related issues regarding results. |  |
| Comments: |  | | | | | | Average Score: |  |

|  |  |
| --- | --- |
| Sub-Criterion Definitions: |  |
| Achievement of Outcomes | Achievement of outcomes is determined via direct statements indicating what outcomes were met and to what degree they were met. |
| Comparison of Current to Previous Results | Comparisons of current results to previous results are possible if a previous report is available. In these comparisons, past results are referenced to show how changes in the programming impact current results. |
| Sharing Results | The process of educators distributing results to individuals for which the results hold meaning (colleagues, upper administration, students, parents, accreditors, the field via presentation and publications). |

**Criterion 8 – Use of Results for Improvement**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Does the report include information about how the results will be used to improve programming or assessment processes? | | | | | | | |
| **If yes, then …** | | Complete current block | | **If no, then …** | | Provide an “average score” of zero, and move to next criterion | |
|  | | | | | | |  |
| **Use of Results** | The primary purpose of assessment is to use the results to make programmatic changes that improve student learning/development when necessary. | | | | | |  |
| **Sub-Criteria** | **Exemplary (3)** | | **Proficient (2)** | | **Developing (1)** | **Missing (0)** | **Score** |
| Programmatic Improvement | Examples of improvements to (or plans to improve) the *program* are **documented** & **directly linked to findings.**  These improvements are **specific** (e.g., where in curriculum they will occur, approximate dates of completing this work).  Or **clear** justification of ***why*** no program changes are necessary given the findings. | | Examples of improvements (or plans to improve) to the *program* are **documented** and **directly related to assessment findings**. However, improvements **lack specificity**. | | Examples of improvements to the *program* are **documented**, but the link between them and the assessment findings is **not clear.** | **No mention** of any *program* improvements.  **No justification** for not improving the program. |  |
| Assessment Improvement | **Critical** **evaluation** of past (if applicable) and current *assessment* *process*, including acknowledgement of flaws; both **past improvements** (if applicable) and **intended improvements** are provided. For both, **specific details** are given.  Or **clear** justification of ***why*** no changes to the assessment process are necessary. | | **Critical evaluation** of past (if applicable) and current *assessment process*, including acknowledgement of flaws; plus, **evidence of some moderate revision** or **general plans** for improvement of assessment process. | | **Some** **critical** **evaluation** of past (if applicable) and current *assessment process*, including acknowledgement of flaws, but **no evidence** of improving upon past assessment or making plans to improve assessment in future iterations. | **No mention** of how this iteration of *assessment* is improved from past administrations (if applicable) or plans to improve future iterations. |  |
| Equity Centered | **Detailed, intentional** **actions** to address systemic barriers to equity are **shared** with a **timeline**. For example, data are used to identify causes of inequitable outcomes and plans to address differential effectiveness of programming across groups are articulated.  Actionable findings serve as an opportunity to advance equity. Plans are explained ***in detail* and align** with information in other sections of the report. | | Actionable findings serve as an opportunity to advance equity. **Possible,** **generally described** plans to address equity concerns are reported in **alignment** with information presented in other sections of the report. | | Actionable findings are addressed as opportunities to advance equity, but **no plans** on how to do so are reported. | **No mention** of how actionable findings can be used to advance equity. |  |
| Comments: |  | | | | | Average Score: |  |

|  |  |
| --- | --- |
| Sub-Criterion Definitions: |  |
| Programmatic Improvement | Programmatic improvement involves proposed or implemented changes to the program itself as a result of the current assessment cycle’s findings. |
| Assessment Improvement | Assessment improvement involves proposed or implemented changes to the assessment process of the program as a result of the current assessment cycle’s findings. |

**General Feedback**

|  |
| --- |
| **If you have general comments regarding the report you are reviewing, please provide them below.** |
|  |

**Glossary**

|  |  |
| --- | --- |
| **Term** | **Definition** |
| **Adherence** | Programming adherence in the implementation fidelity process is a yes/no question: Was the intended programming element delivered? |
| **Context** | Programming context refers to the who, what, where, and when of the program being implemented. That is, the context is all the details of the program (who is facilitating it, for which students, where, when the program is being implemented, etc.). |
| **Core Aspects of Implementation Fidelity** | The core aspects of implementation fidelity include program differentiation, adherence, exposure, quality, and responsiveness. |
| **Desirable Degree** | The desired level of the outcome as a function of the program. Often represented as the amount of growth/change over time, comparison to previous year's data, comparison to educator's standards, or comparison to external criterion like national standards or other institutions. For example, students will gain 0.50 standard deviations from beginning to the end of the program on the specific outcome; after completing the program, students will engage in 5 self-directed volunteering activities; after completing the program, students can apply 3 self-regulation strategies; after completing the program, students can accurately demonstrate 3 safe-sex practices. |
| **Developed Measure** | A measure/instrument/tool specifically created by those assessing the program to determine whether outcomes have been met. |
| **Direct measure vs. Indirect measure** | Direct measures require students to manifest the skills, knowledge, attitudes and behaviors that are of interest (e.g., performance of a task, test scores, endorsement of values, engagement in particular behavior). Indirect measures rely on indirect evidence to infer that students possess certain skills (e.g., students rate their confidence to perform the task, students rate their perceived ability in a content area, students provide ratings of satisfaction). |
| **Effectiveness** | A measure of quality based on the degree to which the program is successful in achieving the stated outcomes. |
| **Etiology** | The study of causes or origins of the distal outcome. What knowledge, attitudes, & skills are needed to achieve the distal outcome for different populations. |
| **Exposure** | Exposure to programming is defined by specifying the planned time for each program element, then recording how much time was allocated/spent. Aids in determining whether students were exposed to the full intended program. |
| **Implementation Fidelity** | The degree to which the programming was delivered as intended. |
| **Implementation Fidelity Data** | Data gathered using an implementation fidelity checklist which is used to evaluate the degree of alignment between the delivered and intended program. Implementation fidelity data are collected alongside outcomes data or before collecting outcomes data. |
| **Improvement** | Improvement is evident when current results can be compared to past results. For example, making a programmatic change cannot be considered an improvement until assessment is completed and compared to previous trends and gains in outcomes are evidenced. |
| **Measure** | A measure is a cognitive test (25-item multiple choice test of intercultural competence), a non-cognitive instrument (self-reported value of diversity), or rubric of observable behavior or products (observations of students' ability to lead a group, produce of a student created resume) which is used to operationalize an outcome. The term "instrument" is interchangeable with the term "measure." |
| **Measure Development Process** | The measure development process refers to the steps that should be taken when one is interested in developing a measure to assess an outcome. |
| **Outcomes** | Outcomes describe what students know, think, value and can do. Outcomes are often stated as *desired* outcomes—outcomes educators *hope* students can achieve as a function of intentional programming. Outcomes are specific; they specify what students will do (e.g., explain, compare, build, engage, commit to). |
| **Population** | The group of students that educators intend to impact through the program (e.g., first-year students, transfer students, international students, alternative spring break students, fraternity or sorority students) |
| **Positionality** | An individual's relative social, cultural, and political location in relation to another person in a particular context. Positionality is closely related to a person’s social identities, standpoints, and cultural practices. |
| **Program Differentiation** | Program Differentiation is a component of the implementation fidelity process. The program features believed to facilitate mastery of each outcome are aligned with the outcomes. |
| **Programming or Program** | "Programming" or "program" is used as an all-encompassing term to include any tactics, strategies, practices, activities, experiences, content, and pedagogy that is intentional to achieve desired student learning and development outcomes (change what students know, think, or do).  Events created with the sole intention of entertaining (e.g., movie theater, easter egg hunt, pizza parties, pool tournaments, concerts, comedy shows) do *not* need to be assessed for increases in student learning and development, as that is not their intention; these "events" are not considered "programs" for the purposes of this document. “Programs” in this document are educational and developmental in nature. |
| **Psychometric Properties** | Indicators of the reliability of scores and validity of inferences from the scores. There are a variety of methods that can be employed to gather reliability and validity information. Measures used without such information are suspect as it is unclear if the scores represent the outcome of interest. |
| **Quality** | Quality ratings in the implementation fidelity process capture how well a program element was delivered or implemented (e.g., clear, confusing, rushed). |
| **Reliability** | The consistency of scores, often represented by one of three types of reliability estimates: internal consistency, inter-rater agreement, and test-retest reliability. |
| **Responsiveness** | Responsiveness in the implementation fidelity process are ratings that capture the degree to which students are engaged or actively participating in the program being assessed. |
| **Stakeholders** | Individuals/groups that have an interest in the program and can affect or be affected by the program. |
| **Student-Centered** | For Student Learning Outcomes to be student-centered, they need to be stated in a way that reflects the knowledge, attitudes, and skills that *students* are expected to gain as a result of participating in the program. **Good Example:** As a result of completing Transfer Student Orientation, incoming transfer students will be able to list 4 academic resources on campus. **Bad Example:** Facilitators will deliver presentation about academic resources. *(Focuses on what facilitator will do, rather than what students will learn.)* |
| **Universal Design for Learning (UDL)** | Universal Design for Learning (UDL) is a framework to improve and optimize teaching and learning for all people based on scientific insights into how humans learn. In this framework, learning experiences should provide multiple means of engaging with content, multiple means of representing the content, and multiple means of action and expression for students to demonstrate learning the content. |
| **Validity** | The degree to which evidence and theory support interpretations of test scores for proposed uses of tests. |

The contents of the *Student Affairs Assessment Improvement Rubric* were informed by the following existing meta-assessment rubrics:

* University of the District of Columbia (2020). “Meta-Rubric for Evaluating Assessment Plans and Reports.”

<https://docs.udc.edu/assessment/Meta-Rubric-12-10-20.pdf>

* James Madison University (2015). “Assessment Progress Template (APT).”

<https://www.jmu.edu/assessment/AcademicProgram/AssessmentReporting.shtml>

* Andrews University (2015). “Rubric for Evaluating Program Assessment Plan.” [https://www.andrews.edu/services/effectiveness/assessment/resources/andrews-rubric-for-evaluating-program-assessment-plans-and-reports-](https://nam10.safelinks.protection.outlook.com/?url=https%3A%2F%2Fwww.andrews.edu%2Fservices%2Feffectiveness%2Fassessment%2Fresources%2Fandrews-rubric-for-evaluating-program-assessment-plans-and-reports-v.2.docx&data=05%7C02%7Cfinneysj%40jmu.edu%7C6a48ba98328a4467562a08dc264d5018%7Ce9333c23cac742f499895cee3d4a79c0%7C0%7C0%7C638427360982589780%7CUnknown%7CTWFpbGZsb3d8eyJWIjoiMC4wLjAwMDAiLCJQIjoiV2luMzIiLCJBTiI6Ik1haWwiLCJXVCI6Mn0%3D%7C0%7C%7C%7C&sdata=9aMaWb7FpUl9%2F7FMdH661AtFKrMtUe3FjoEMj3XBp2Q%3D&reserved=0) [v.2.docx](https://nam10.safelinks.protection.outlook.com/?url=https%3A%2F%2Fwww.andrews.edu%2Fservices%2Feffectiveness%2Fassessment%2Fresources%2Fandrews-rubric-for-evaluating-program-assessment-plans-and-reports-v.2.docx&data=05%7C02%7Cfinneysj%40jmu.edu%7C6a48ba98328a4467562a08dc264d5018%7Ce9333c23cac742f499895cee3d4a79c0%7C0%7C0%7C638427360982597792%7CUnknown%7CTWFpbGZsb3d8eyJWIjoiMC4wLjAwMDAiLCJQIjoiV2luMzIiLCJBTiI6Ik1haWwiLCJXVCI6Mn0%3D%7C0%7C%7C%7C&sdata=uZTnOOWS3oj9lZ0fAGsNus5B8HF2k3tmG8gAQFIcI3U%3D&reserved=0)