

Standards Overview by Categories

U Utility Standards. The utility standards are intended to ensure that an evaluation will serve the information needs of intended users.				
	RELEVANCE TO CULTURAL COMPETENCY	STRENGTHS AS CURRENTLY WRITTEN	CONCERNS/ LIMITATIONS	RECOMMENDATIONS
Standards.	The utility, value and influence of evaluation is implicitly defined in cultural context. Issues of evaluation credibility & influence & who is competent to perform evaluation are tied to cultural context.	Standards are written at a general level & can be called upon in making arguments regarding cultural competence.	Context (U5) not explicitly inclusive of culture. U6 unnecessarily restrictive to intended users within immediate time frame.	Standards are mostly appropriately written & should be retained. U4 should include standpoints or culturally defined perspectives in list of descriptors. Revise U5 to "...provided to and understood by multiple stakeholder audiences." Retain standard on dissemination issues but make separate standard re issues of time & timing (U6).
Overviews.	Overviews touch on relevant issues of values, power, voice & entitlement. Culture is named in U6.	Reference made to diverse stakeholders (U1) & need to promote inclusion. Good attention to overriding legal, moral & ethical concerns (U3). U4 reflects core of cultural competence in values. U6 addresses issues of entitlement in access to reports, language translation & need to tailor report to fit culture.	Stakeholders for U1 appear to give priority to management perspectives. As written, attention to less powerful groups is marginalized as a "special" concern vs. a core concern in all evaluation. Cultural awareness, sensitivity not noted as part of evaluator credibility (U2). Also, credibility goes beyond ability to defend. As written, fails to address culturally-relevant issues such as historical legacy, respect. "Evaluator preferences" understates the role of culture and position in shaping values (U3). Issues of power surrounding values not addressed (U4). Also, U4 leaves false impression that values are largely about interpreting findings vs. infusion of values in entire process. Positions audience as passive recipients of reports (U5). Issues of time & timing get lost in discussion of dissemination (U6). Tailoring report to cultural context not well developed. U7 limited by narrow definition of use as intended, results-based & immediate, instrumental + conceptual. Description of evaluation role condescends to program persons.	Elaborate consumer perspectives (U1), expand examples of cultural diversity to include economic status, ethnicity, sexual orientation, education, age, disability, religion, gender, health status, immigration status to illustrate breadth. Include cultural competence in U2 characteristics & add credibility with stakeholders. Tie credibility to integrity. Add equity, fairness & social justice to U3 list of overriding concerns. Strengthen language about cultural position & values of evaluator. Add cultural relevance to illustrative list of concerns about both program and evaluation. U4 could benefit from stronger statement regarding centrality of values identification & whose perspectives are included /omitted to the entire evaluation process, not just findings interpretation. Expand U5 to include non-written communication & include dialog with audiences regarding report clarity & meaning. Expand U6 to develop discussion of cultural context. Revise U7 to broaden definition of impact to include unintended (negative + positive), process-based & longer-term influences. Describe roles of program personnel & evaluator in more respectful, collaborative terms.

Standards Overview by Categories

U Utility Standards. The utility standards are intended to ensure that an evaluation will serve the information needs of intended users.				
	RELEVANCE TO CULTURAL COMPETENCY	STRENGTHS AS CURRENTLY WRITTEN	CONCERNS/ LIMITATIONS	RECOMMENDATIONS
Guidelines.	Guidelines support attention to non-discrimination, respect for cultural differences in communication, considerations of power differentials among cultural subgroups, and reflection on treatment of culture in designing evaluation.	U1.G explicitly addresses nondiscrimination on basis of gender, ethnicity or language background. U1.B raises important issue of who represents stakeholders. Attention to social/political forces (U2.A) specifically cites cultural differences. Tailoring reports to audiences (U5) is culturally congruent. U5.H-I address important cultural issues. U7.E is written to communicate greater balance, respect for stakeholders than other U7 guidelines.	Appropriateness for cultural context should be included in review of evaluation plan (U2.D). Credentials for reviewers (U2.C) should include cultural competence. U4 gives no explicit attention to culture. Formulaic guidelines may not yield culturally-relevant questions. Too much specificity in guidelines may limit utility across contexts—e.g., U3.H might work against allocating sufficient time & resources to answering culturally relevant questions well. Answering such questions well may be both labor-intensive and time-consuming. U6 presumes linear, monochronic time frame, preordinate design & a written report. U6.J frames diversity as a “social impediment.” Very offensive, inappropriate. U7 privileges evaluator as person with answers & stakeholders as persons to be convinced—power differential permeates several guidelines. U7.G assumes linear, results-based, unidirectional influence.	Include both formal & informal leadership roles in U1.A. Expand diversity dimensions relevant to nondiscrimination in U1.G & to social/political forces in U2.A. Strengthen language to incorporate cultural context in guidelines. Add, “understand the cultural context of the evaluation” to U3.A. Address complexity of who speaks for stakeholder groups in U4, along with examples that illustrate cultural congruence (U4.C). Stress inclusion of multiple perspectives. Add a U5 guideline that makes it clear that cultural dimension should be considered in determining most appropriate media or other aspects of reporting process. Discussions of translation should include ASL & Braille. Remove/revise U6 to focus on <i>strengths</i> of cultural traditions in choosing strategies & timing. Revise U7 guidelines to remove power differential, expand conceptualization of impact/influence.
Common Errors	Omitting attention to culture can compromise evaluation. Utility errors point to omissions as well as errors of action & interpretation that undermine cultural competence.	U3.B addresses the importance of multiple voices. U6.A&G have potentially important implications for cultural competence as they both address inclusion. U7.A cites stakeholder disrespect as an error & U7.C supports consideration of stakeholder values.	Failing to include evaluator’s cultural competence as a dimension of credibility (U2.A). Student status should not be automatically equated with inexperience (U2.I). Not clear where advocacy models of evaluation fit with regard to U2.B. When too prescriptive, Guidelines & Common Errors may compromise cultural competence. Cultural aspects of values identification not made explicit. Elitist language in U5 privileges technical	Add as a U1 error “Failing to anticipate competing or adversarial views of program goals & objectives held by stakeholder groups.” Add cultural competence to list of credibility concerns. Delete “student” from U2.I. Errors cited should not be biased against advocacy evaluations. Be sure that labor-intensive strategies to establish cultural credibility aren’t cast as errors. Revise U2.D to include both setting & cultural context. Add “and cultural perspectives” to U3.A statement on multiple voices. Instead of specifying <i>how</i> information

Standards Overview by Categories

U Utility Standards. The utility standards are intended to ensure that an evaluation will serve the information needs of intended users.				
	RELEVANCE TO CULTURAL COMPETENCY	STRENGTHS AS CURRENTLY WRITTEN	CONCERNS/ LIMITATIONS	RECOMMENDATIONS
			written reports, implying a false dichotomy between precision & clarity.	requirements should be updated, note the importance of follow-up procedures that are congruent with & respectful of norms of each stakeholder group. Add “cultural perspectives” to U4.B plus a new U4 Error, “Failing to educate oneself in the cultural significance of decision rules and rituals.” Add culture-based errors to U5—e.g., “Failing to consider cultural variables that define appropriate and effective communication when deciding how to report information;” “Assuming that English is necessarily the appropriate language in which to communicate and that technical terms understood by the evaluator are familiar to the audience.” Add “& its cultural diversity” to “describe the program” in U5.H. Add civil rights, social justice & equity issues to list of overriding concerns in U6.E. Remove “target” from U7.G; restate in a way that is more respectful, e.g., “Failing to attend to the needs of specific stakeholder audiences.”
Illustrative Case Descriptions	Cases offer strong potential to explore the nature of influence at macro/mezzo/macro levels & reveal culturally embedded nature of values. As written, most Utility case scenarios do not realize this potential.	U4.1 explicitly mentions race, urban location (albeit from a deficit perspective). U4.2 attends to age diversity, rural location. U7.1 introduces issues of power & authority & provides opportunity to examine congruence of evaluation with values of program being evaluated.	No mention of cultural context in U1.1, which also fails to address program participants as stakeholders. Cultural context information also missing from U2.2, U3.1, U5.1-2, U6.1-2 & U7.1-2. Economic status is relevant to U1.2 but not addressed. U2.1 is fraught with institutional racism. Assumes “minority” refers to race. U4.1 takes deficit perspective in dealing with race, Ebonics. U5 cases are skeletal & similar & neither brings out cultural dimensions well.	Infuse cultural diversity in case illustrations in relevant ways, avoiding token references. Replace U1.2 with stronger illustration of standard. U2.1 is unacceptable as written. It should either be replaced entirely or the analysis rewritten to point out how judgments of credibility are infused with values from context, including discriminatory ones. Add a 2 nd case to U3 that draws out cultural dimensions of standard more clearly. Consider a separate standard on time that goes beyond reporting issues. Replace “handicapped” in referring to persons with disabilities. Try to frame analyses to raise questions & suggest possibilities rather than prescribing a single course of action (U6.1). Replace U6 cases with ones that better illustrate power dynamics.

Standards Overview by Categories

U Utility Standards. The utility standards are intended to ensure that an evaluation will serve the information needs of intended users.				
	RELEVANCE TO CULTURAL COMPETENCY	STRENGTHS AS CURRENTLY WRITTEN	CONCERNS/ LIMITATIONS	RECOMMENDATIONS
<i>Illustrative Case Analyses</i>	Analyses hold potential to examine the cultural context of evaluation influence and of inclusion that shapes the evaluation process. Position of program participants/consumers in relation to the evaluation is especially relevant to Utility standards. Potentially examines power and influence.	Analysis U1.1 challenges authority-driven definition of stakeholders, emphasizes more balanced perspective. Analysis U4.1 hints at diversity <i>within</i> standpoint perspectives.	U2.1 fails to identify or challenge racism visible in the scenario or take up issues of systemic bias embedded in the judgments of credibility. U4.1 similarly misses opportunity to examine how societal values/prejudices get drawn into evaluation. U4.2 analysis fails to catch omission of older adults from planning process. U5.2 fails to draw attention to matching communicative strategies to culture of setting even though suggestions made illustrate this with respect to organizational culture. Prescriptive tone of analysis in U6 suggests only one right way to proceed. Lack of cultural information in U7.1 hinders analysis of why parents deferred to evaluator & withdrew. Issues of power & authority not addressed. Participant perspectives not evident in U7.2 nor are they queried in the analysis.	Rewrite the analysis of U2.1 to challenge institutional racism evident in this case. Analyses should note position of program participants as stakeholders & challenge systemic bias (e.g., along lines of race, economics, age, disability) in scenarios.

Standards Overview by Categories

F Feasibility Standards. The feasibility standards are intended to ensure that an evaluation will be realistic, prudent, diplomatic, and frugal.				
	RELEVANCE TO CULTURAL COMPETENCY	STRENGTHS AS CURRENTLY WRITTEN	CONCERNS/ LIMITATIONS	RECOMMENDATIONS
Standards	Feasibility standards support entering a context respectfully, minimizing intrusion, & giving something back	These standards give specific attention to politics & speak to inclusion & diversity of stakeholder views	Cultural lines along which politics and other differences break are not addressed. Issues of power not well articulated. F2 positions diversity as a negative influence to be counteracted. F3 is simplistic in its treatment of costs/benefits of evaluation.	Retain standards with revised wording that is more concise (F2), grammatically correct (F3) & does not put stakeholders in a one-down position. In rewording F2, make sure that it does not take a deficit approach to diversity; benefits of collaboration to achieve political viability merit equal attention.
Overviews	Politics and values are addressed directly. Issues of diverse stakeholder values are particularly relevant to cultural diversity.	Touches on issues of fairness & equity in F2, which also addresses power more directly than other standards.	Gives only limited attention to fairness & equity. Positions evaluators as the “good guys.” Motives of others become suspect. Oversimplifies the synthesis if viewpoints in analysis of costs, effects.	Give greater attention to respect & giving back to the program/community as part of concern for feasibility. Avoid implicit assumption that evaluator is in the right & others either are to be convinced or defended against.
Guidelines	Several guidelines support inclusion of diverse voices and implicitly suggest respect for context. Also the fact that the costs & benefits of evaluation may be perceived differently is important for cultural consideration.	Specifies a number of reasonable procedures that could co-exist with cultural competence. Gives reasonable attention to multiple perspectives, in general sense, not specific to culture.	Cultural competence not included as a necessary qualification. Issues of power/authority understated. Procedures supporting cultural competency are often labor intensive & time-consuming; they might be omitted under these guidelines.	Include guidelines that address need for long-term relationships to achieve cultural congruence in some settings. Avoid assumption that all designs are preordained. Make clear that cultural competence requires an investment; it should not be dismissed as “inefficient” or infeasible. Wording changes would strengthen specific guidelines.
Common Errors	Errors in Feasibility may lead to evaluations that are not culturally responsive, yield little benefit & may cause harm.	Directs attention to context. Acknowledges organizational power structure.	Potentially creates false dichotomy of practicality versus accuracy. Doesn’t address culture in discussions of context, stakeholders, or power.	Be more explicit in noting lack of attention to cultural context as a common error. Cite failure to allocate sufficient time to enter respectfully & build relationships of trust as common errors.
Illustrative Case Descriptions	Though not written that way now, the cases could potentially include context information and illustrate how culture was addressed in entering a community, working together smoothly, & handling differences.	Some attention given to economic status, age, and location (rural, suburban, urban)	Examples focus on design issues & details, with little acknowledgement of context, little attention to power dynamics. Also content of F1.1 is very dated. F3.1 is quite simplistic, of limited value.	Cases F1.1 & F3.1 should be replaced. Cases should illustrate more than feasibility of implementing certain designs. Case F3.1 should bring in more complex issues of evaluation’s costs (money, non-money, opportunity) and benefits, including attention to cultural position of evaluation. Greater attention to cultural diversity in describing context of cases.

Standards Overview by Categories

F Feasibility Standards. The feasibility standards are intended to ensure that an evaluation will be realistic, prudent, diplomatic, and frugal.				
	RELEVANCE TO CULTURAL COMPETENCY	STRENGTHS AS CURRENTLY WRITTEN	CONCERNS/ LIMITATIONS	RECOMMENDATIONS
<i>Illustrative Case Analyses</i>	Potentially, the analyses could include an examination of cultural dynamics or raise possible questions. (They are not written that way now.)	Analysis F1.1 does address the complexity & politics of educational systems.	No reflection on cultural context, even when noting factors such as mobility & attrition that could be culturally relevant. Analyses avoid complex issues of values, tend toward reductionism, oversimplification.	Analyses should raise questions about power relationships, political context. More thoughtful consideration of complexities needed. Too simplistic as written.

Standards Overview by Categories

P Propriety Standards. The propriety standards are intended to ensure that an evaluation will be conducted legally, ethically, and with due regard for the welfare of those involved in the evaluation, as well as those affected by the results.				
	RELEVANCE TO CULTURAL COMPETENCY	STRENGTHS AS CURRENTLY WRITTEN	CONCERNS/ LIMITATIONS	RECOMMENDATIONS
Standards	Propriety standards capture issues at the core of cultural competence: equity & access (P1), respect (P4) + dynamics of privilege & position. Justice & fairness (P5), inclusion (P6), honesty, ethics (P7), accountability (P8).	These standards as written create an extremely important platform for examining issues of ethics, accountability, & civil rights. Resources are broadly defined to include human capital (P8). Nothing is incongruent with consideration of culture.	Some of the language is restrictive & alienating—e.g., “targeted” participants (P1) & “human subjects” (P3). P2 presumes formal agreements are written, which may not apply to all cultures.	Delete “targeted” from P1. In P3, replace “human subjects” with “participants in the evaluation process.” P8 title should consider a change to be clear that all resources—money & non-money—should be included.
Overviews	Overviews could make the cultural intersections visible. As written, they introduce important concepts such as fairness & call for self-scrutiny in matters such as conflict of interest. Reflection & self-awareness are key to cultural competence.	P1 & P4 broaden persons affected by evaluation beyond intended direct consumers to community & society. P2 calls for mutual respect. Good that P3 seeks to protect persons who are receiving goods & services. P6 takes strong advocacy stance that is congruent with inclusion & supports multicultural validity. P7 raises important points regarding the infusion of bias in both internal & external evaluation.	Cultural differences in negotiation &/or recording agreements not acknowledged. P2 presumes a preordinate design. P4 attention to “hurt feelings” risks trivializing important issues of understanding & respect across cultural differences. The deficit perspective of P5¶2 should be challenged. P7 is skewed toward financial conflicts of interest. P8 narrows & restricts the standard toward fiscal audits & accounting, ignoring non-money resources such as cultural capital, responsible <i>pro bono</i> work.	P1 overview could benefit from elaboration regarding conceptualization of public good & diverse perspectives of managers, providers & direct + indirect consumers. P2 should be rewritten to acknowledge cultural differences in negotiating agreements. Add a ¶ to P3 to give respect equal emphasis with legal rights & protections. Address respect at both micro & macro levels. Note cultural intersections explicitly in the overview. Going in with avowed purpose of determining weakness (P5¶2) is not culturally competent; reword. Retain the strong stance of P6 toward disclosure, utility & fairness. Balance financial & personal conflicts of interest in P7. P8 needs to be expanded to include responsible use of resources in <i>pro bono</i> or barter contexts.
Guidelines	Culturally competent practice is supported by clear operationalization of “cultural sensitivity.” Best examples are found in P3 & P4.	P3 explicitly mentions cultural values! Also addresses civil rights & communication with language diversity. P4 addresses culture, social values & language differences. P5 addresses issues of thoroughness, completeness & inclusion of diverse perspectives, which could include cultural diversity. P6	Guidelines for P1 seem to limit its scope, reversing the positive tone of the overview. Also not privileged position—informing stakeholders vs. working with them to clarify purpose of evaluation. No attention to cultural competence in the agreement (P2). Include tribal laws when listing legal examples. P3 guidelines use disrespectful language, “human subjects” & “language minority.” P3	Reword & reorder Guidelines + add new content. Add a guideline on respectful reporting of negative findings. Rewrite P1 guidelines, adding a guideline on equity (e.g., “Evaluators should consider the cultural dimensions relevant to the participants and the public in the program context and examine the delivery and effectiveness of services in meeting the needs of these persons across dimensions of difference.”) Develop a separate set of guidelines for emergent designs or rewrite these to bridge

Standards Overview by Categories

P Propriety Standards. The propriety standards are intended to ensure that an evaluation will be conducted legally, ethically, and with due regard for the welfare of those involved in the evaluation, as well as those affected by the results.				
	RELEVANCE TO CULTURAL COMPETENCY	STRENGTHS AS CURRENTLY WRITTEN	CONCERNS/ LIMITATIONS	RECOMMENDATIONS
		can support inclusion of cultural context & justification for including or excluding cultural dimensions. P7.F explicitly names culturally defined advantages—social, moral & political—broadening this standard in important ways.	mentions organizational context but not community or culture. By operationalizing P5 in terms of reporting issues, guidelines undermine the scope of impact of this standard. Culturally competent communication not simply a matter of linguistic form (P6). Cultural dimensions of “different perspectives” are not explicit (P7+P5). P7 is too directive in E&G. Emphasis on frugality may undermine cultural competence (P8). Non-money costs (e.g., time of volunteers) should be included; also include any persons serving as cultural guides who are not paid for their time.	designs evenly. Don’t continue to privilege preordained designs & tell emergent to “adjust.” Rewrite guidelines to reflect culturally appropriate systems of governance, beyond a “management” orientation. Revise language of P3.F & O; delete N. Add a P3 guideline that directs evaluators to examine their <i>own</i> values & assumptions as well as participants’. Add depth, detail to P4, a potentially critical standard in multicultural work that is underdeveloped. Expand discussion of human dignity & worthy & countervailing issues of prejudice, discrimination & disrespect. Remove qualifying language—e.g., “make every effort” to make statement more assertive. In addition to organizational context, familiarity with community & cultural contexts. Expand P5 to include attention to thoroughness, fairness & diversity in designing & implementing the evaluation, not just reporting results. Revise P7.G&E to direct evaluators to consider implications of the choices of pathways & authority rather than unilaterally supporting one mechanism. Expand P8 to address non-money costs & cost estimates under emergent designs.
Common Errors	Failures to consider needs of consumers, community & public-at-large compromise cultural competence. Cultural competence could also be an important contractual element. Common Errors prompt one to reflect on sources of bias including cultural sources.	Recognizes collaboration in negotiating agreement (P2). Addresses protections for children & “language minority” participants (P3) & age, sex, ethnicity & cultural background (P4). Promotes nondiscrimination. Power dynamics touched on in P4.A. P4.E addresses personal attributes that are culturally defined. Refers to multiple perspectives (P5) that shape	P1.H labels participant/community advocacy as an error. Time frame & what constitutes common sense (P2) may be culturally defined. Limited attention to diversity beyond age & language. P4.B would again seem to argue against affirmative action to include perspectives of vulnerable populations. Language choice sometimes trivializes concerns. P5.C&E are operationalized primarily in terms of reporting results (vs. issues permeating the entire	Strengthen the language upholding P1. As written, Common Errors undermine the standard itself. Reword P3 to make clear that the general error is in overstating the boundaries of protection that evaluators can promise/control. Shift language way from human subjects & address failures to recognize context-relevant risks more clearly. Address full range of diversity characteristics, vulnerabilities. Make language stronger regarding violations, moving beyond “embarrassment” or “hurt feelings.” Broaden P5 beyond reporting concerns (a holdover from 1 st ed.). Strengthen language of

Standards Overview by Categories

P Propriety Standards. The propriety standards are intended to ensure that an evaluation will be conducted legally, ethically, and with due regard for the welfare of those involved in the evaluation, as well as those affected by the results.				
	RELEVANCE TO CULTURAL COMPETENCY	STRENGTHS AS CURRENTLY WRITTEN	CONCERNS/ LIMITATIONS	RECOMMENDATIONS
		interpretation. P7.D can include persons with unique cultural qualifications.	process). Underestimates extent to which all interpretations & actions are shaped by cultural perspectives. Weak wording suggests standards may be optional social niceties vs. mandates. Should note that evaluation procedures “well established” in majority context may introduce bias in minority contexts (P7). Seniority may not reflect highest qualification with respect to cultural variables.	P5.J. Move up P5.K to give greater prominence. Add as error “sacrificing cultural competence to save money.” In judging who is qualified (P8) make sure that cultural competence has been taken into account.
<i>Illustrative Case Descriptions</i>	Cases potentially rich in illustrating issues of propriety & ethics rooted in cultural difference.	P4.1 addresses economic diversity; poverty-level status of students is relevant to the disrespect shown them. P4.2 illustrates sexism within a large corporate environment. P6.1 gives explicit attention to cultural content in desegregation as evaluand.	As in other sections, cultural information is substantially absent, implying that cultural context is irrelevant to proper use of the standards. No cultural information provided in P2.1-2, P3.1-2, P5.1-2, P6.2, P7.1-2, P8. P1.1 places evaluand in an urban school but with no other details. P4.1 labels students as “poverty-level” but offers no additional context information. P5.1 doesn’t illustrate issues of fairness. P6.1 speaks of majority & minority groups with no information on community context or diversity within groups. Also no information on evaluator characteristics. Links to P7-8 lost in other case information.	Add two contrasting case descriptions to P1 & remove the current weak illustration. Avoid representing diversity with single labels. Token inclusion of cultural markers sets poor example. Include information on evaluator characteristics. Use positive case illustrations to show how challenges were met rather than to present a case with few challenges. Add a second case to P8 that goes beyond financial budgeting OR add a separate standard to address broader issues of responsible documentation & oversight of non-money resources & opportunity costs.
<i>Illustrative Case Analyses</i>	Analyses potentially shed light on how culture intersects judgments of proper professional behavior (see P4.2 for an analysis that takes positive steps to do this).	P2.2 includes discussion of trust. P4.1-2 explicitly address cultural diversity. P4.2 addresses disrespect shown in stakeholder exclusion. P6.1 stresses need to seek both majority & minority input & responses to report. Also	P1.1 suggests strategies that exceed allotted resources & shift evaluand. P2.2 suggests strategies that may be incongruent with some organizational cultures. Analyst suggests a design that’s not responsive to the stated purpose of the evaluation (P5.1). Analyst	Edit analyses more carefully. Make sure that analyses reflect knowledge of regulations regarding privacy, rights & protections—HIPAA, ADA, etc. Address issues of cultural competence in analyses—both at systems and individual levels.

Standards Overview by Categories

P Propriety Standards. The propriety standards are intended to ensure that an evaluation will be conducted legally, ethically, and with due regard for the welfare of those involved in the evaluation, as well as those affected by the results.				
	RELEVANCE TO CULTURAL COMPETENCY	STRENGTHS AS CURRENTLY WRITTEN	CONCERNS/ LIMITATIONS	RECOMMENDATIONS
		addresses compliance with federal & state laws.	recommendations (P7.2) may be questionable fit with organizational culture. Criticism is mild in light of violations (P3.1). P4.1 analysis does not point out the evaluators' lack of cultural competence regarding the organizational culture of secondary school or culture of poverty. P4.2 fails to address systemic issues of organizational culture, power, authority, status, nor does it point to sexism evident in this scenario. Completeness is equated with fairness (P5.2); broader issues of fairness not addressed. Issues of racism not addressed (P6.1) nor are dynamics of power at community, societal levels. Lacking cultural information, impossible to determine if problems cited are culturally connected (P7.1), gender related (P8).	

Standards Overview by Categories

A Accuracy Standards. The accuracy standards are intended to ensure that an evaluation will reveal and convey technically adequate information about the features that determine worth or merit of the program being evaluated.				
	RELEVANCE TO CULTURAL COMPETENCY	STRENGTHS AS CURRENTLY WRITTEN	CONCERNS/ LIMITATIONS	RECOMMENDATIONS
Standards.	These standards speak to validity & rigor. Cultural considerations must not be seen as antithetical to rigor. Appropriate consideration of culture supports rigor.	Standards support good evaluation & metaevaluation, setting forth relevant parameters for examining culture.	Validity (A5) & reliability (A6) are narrowly construed in terms of measurement issues. Many missed opportunities to address cultural context (A2) & to address macro issues such as white privilege (A11).	Expand A2 to make cultural context explicit, pointing out the bi-directional nature of the culture/evaluation relationship. A7 should include review for cultural bias. Expand A11 to include macro issues and A12 to include cultural critique. A8-9 are clearly written but may perpetuate a dichotomy of analysis & interpretation that's not helpful.
Overviews.	Overviews provide opportunities to recognize cultural connections with accuracy. Can potentially address a range of demographic diversity variables at micro-mezzo-macro levels. Metaevaluation supports multicultural validity.	Good connections to validity though cultural meanings not explicitly addressed (A3, A5). Age diversity addressed (A6, A8) & SES in A8. Attention to emergent designs (A9) is congruent with culturally responsive models.	Culture cited only as a deficit concern (A2). A3 omits culture as a dimension of documentation & monitoring. A4 is particularly weak, conflating information sources with data collection strategies. Overall, quantitative methods are privileged over qualitative, preordained designs over emergent, and micro over macro perspectives. A12 devalues formative, internal metaevaluation, yet cultural competence demands such reflection.	Make cultural connections more explicit in discussing each standard, correcting errors of omission. Expand discussions of validity to include critical multiplism & interactions of validity & culture. Quality control review should monitor coverage of both majority and minority standpoints. Quantitative analysis should include theory-based attention to demographic subgroups, examining both similarities and differences. Expand A9 to illustrate other types of analyses such as those based in standpoint perspectives—e.g., queer theory. Avoid implicit privileging of quantitative over qualitative methods. Consider systemic sources of bias that challenge impartiality (e.g., societal attitudes toward age, race, disability). Make explicit the benefits of metaevaluation in supporting cultural competence & multicultural validity.
Guidelines.	Guidelines offer potential action steps for checking cultural congruence of “technical” dimensions. Some make connection explicit (A5); others (A7, A9, A10) create a framework for examining culture but don't take it up explicitly.	Importance of multiple perspectives emphasized (A1). A5 includes clear statement of connection between context, personal characteristics, & validity. Link to culture explicit though not fully developed. Evaluator reflection (A6) can include cultural values, positions. Good connection of justification and validity issues	A1 does not address how multiple perspectives are to be integrated. Cultural & historical context not mentioned in A2. Guidelines often presume an experimental paradigm or preordained design. Unintended influence and process-based influence are not well addressed. Reports are presumed to be written, not always the most culturally-appropriate choice. A4 conflates	Include cultural context in evaluand documentation. Expand A2 to include cultural & historical context plus issues of power & relationship. Guidelines should reflect broad understanding of use/influence in current theory. Include culturally appropriate communications that extend beyond written reports. A4 should address the culturally-bound nature of information sources as a dimension of adequacy. Consider separate standards on sampling and data gathering strategies so that cultural

Standards Overview by Categories

A Accuracy Standards. The accuracy standards are intended to ensure that an evaluation will reveal and convey technically adequate information about the features that determine worth or merit of the program being evaluated.				
	RELEVANCE TO CULTURAL COMPETENCY	STRENGTHS AS CURRENTLY WRITTEN	CONCERNS/ LIMITATIONS	RECOMMENDATIONS
		(A10).	sources of information with issues of sampling & data collection. Cultural competence of metaevaluator not addressed. A12 is imbalanced toward external, summative metaevaluation.	dimensions can be separately scrutinized. Replace “physical handicaps” with disability and expand list of diversity characteristics impacting validity (A5). Expand discussions of practical significance (A8) and cultural context of interpretation (A8-9). Pay attention to epistemological diversity, which intersects cultural diversity. Make sure guidelines address iterative and emergent use of standards (e.g., A11) not just preordinate check off. Delete 12E. Expand 12H into several guidelines so that culture can be explicitly connected to congruence of design with context, culturally appropriate measurement and data collection strategies, culturally-bound interpretations, etc. Avoid guidelines that are procedurally prescriptive when such specifics would vary with context & culture.
Common Errors	Many errors intersect cultural dimensions; a number of these could be made more explicit. For example, “common misinterpretations & inappropriate inferences” (A10) can include misuse of race as an explanatory variable.	Errors create framework within which culture could be addressed (e.g., A1.F could include assuming that program is appropriately implemented across dimensions of cultural diversity). A number of excellent cautions can readily be connected to culture (A5). Good attention given to evaluator’s inability to step outside own cultural context, experience (A6); practical as well as statistical significance (A8); need to examine validity of preconceptions (A9). Acknowledges importance of alternative views of reality & multiple perspectives (A9), which may be grounded in	Because of assumptions that design is preordinate, the need to revisit emerging issues is often lost in the enumeration of errors—especially relevant to emerging cultural appreciations. Prescriptions regarding time may constrain multicultural validity. A3 omits attention to how the purpose of the evaluation gets translated into evaluation questions & whose values are reflected in this translation. A4 does not address the culturally bound nature of what sources are considered “defensible” and how they are defended. Treatment of diversity is limited to “special” or “handicapping” conditions (A5). Somewhat elitist, top-down view implied as to who	Make errors associated with culture more explicit. For example (A1), “Assuming that the majority description of the evaluand represents how it is understood from all cultural perspectives.” Be explicit that ignoring culture is an error. A2 should be expanded to include errors such as taking majority perspective as truth, failing to grasp historical context of evaluand, and failure to understand diversity <i>within</i> culturally diverse subgroups. Make a separate standard on sampling so that issue of cultural diversity can be explored/discussed. Expand areas of cultural diversity addressed in A5 and suggest more inclusive participation—e.g., in reviewing instruments. Note that cultural congruence may require deviation from “standardized procedures.” Be aware of issues of power & position. Avoid guidelines that position the respondent passively in a process

Standards Overview by Categories

A Accuracy Standards. The accuracy standards are intended to ensure that an evaluation will reveal and convey technically adequate information about the features that determine worth or merit of the program being evaluated.				
	RELEVANCE TO CULTURAL COMPETENCY	STRENGTHS AS CURRENTLY WRITTEN	CONCERNS/ LIMITATIONS	RECOMMENDATIONS
		culture, and of unintended outcomes (A10). Early attention to metaevaluation supports critical reflection on culture (A12).	would be qualified to review instruments, procedures. Privileges statistical significance over practical significance (A8). Creates a false dichotomy between rigor & relevance (A8-9). Casts distortion as personal failure rather than endemic to a system or society. Emphasis on authority (A11) may run counter to culturally responsive models. Separating quantitative and qualitative analyses (A8-9) may perpetuate false dichotomies; many Guidelines and Errors overlap.	directed by the evaluator (A7). Recast concept of rigor to include cultural relevance/congruence & update to reflect mixed-method evaluation. Need to acknowledge culturally bound nature of what is considered sound, sufficient information; limitations of perspective should be noted alongside limitations of data & procedures. (A10). In discussing evaluator authority/position, avoid casting participatory practices as error. Tie metaevaluation to critically reflective practice, inclusive of ongoing cultural critique.
<i>Illustrative Case Descriptions</i>	Analyses can call attention to culturally bound assumptions as well as how evaluation is shaped by background & values of evaluator. As written, lack of detail leads to missed opportunities to examine culture.	Cases are rich with opportunity to explore cultural context (A2) though none is cultivated. Some attention to economic diversity (A5), gender (A8), level of education (A8) & urban/suburban/rural location (A10). A1 avoids labeling students seeking help & discusses political & organizational constraints. A12 contains potential for exploring politics & organizational relationships.	No cultural context information given in many of the Accuracy cases (A1.1-2, A6.1-2, A7.1-2, A8.1, A9.1, A12). A1 cases fail to capture complexity of synthesizing multiple perspectives into a coherent description. A2.1 gives very limited historical information, background & context. Only size of school is given in terms of appreciating culture. A2.2 offers no information on culture of corporate organization or of participants. A3 is narrowly operationalized and blatantly sexist; a female evaluator is blamed for the actions of a male administrator who violated the initial agreement. A4 describes evaluand as controversial but gives no cultural context information regarding program, school or district. Insufficient information to reflect on perspectives included and excluded as information sources. A5.2 identifies persons	Develop cases that illustrate culturally relevant <i>guidelines</i> (as opposed to making only general background references to culture). Replace A1.2 with a more timely example, better illustrating evolving nature of program descriptions. Develop cases with more detail of culture (both societal & organization—“Big C & little c”) for A2. A3 raises unchallenged sexism & should be replaced. A4 misses the main point of the standard & should be replaced. (It would work better as an illustration of F2.) Revise A10.1—both case & analysis—to give value positions of evaluator same level of scrutiny as those of program personnel. Replace or revise A11.2 to eliminate bias against formative participatory evaluation. Introduce greater complexity into the cases. Present cases in a way that lays out challenges & contradictions—e.g., inverse relationship between multivariate statistics & policymakers’ use of data (A8.2). Avoid single “labeling” treatment of diversity variables, absent theory or discussion. When issues of time, timing, & time constraints are discussed, give greater attention to cultural dimensions of

Standards Overview by Categories

A Accuracy Standards. The accuracy standards are intended to ensure that an evaluation will reveal and convey technically adequate information about the features that determine worth or merit of the program being evaluated.				
	RELEVANCE TO CULTURAL COMPETENCY	STRENGTHS AS CURRENTLY WRITTEN	CONCERNS/ LIMITATIONS	RECOMMENDATIONS
			<p>living in poverty but offers no opportunity to consider diversity among the population, historical context, etc. Unequal attention to staff & consumer characteristics, though both participated in focus groups (A5.2). Time constraints noted without noting potential cultural relevance of time as a variable.</p> <p>A11.2 appears biased against formative participatory evaluation, which would include a number of culturally responsive models. A12 fails to develop potential for exploring politics & organizational relationships. Analysis focuses on method, procedures, & psychometric properties without grounding in cultural context. Analyses take a deficit approach, looking for problems.</p>	time as a variable.
<i>Illustrative Case Analyses</i>	Cultural analysis hampered by lack of information.	Little direct critique of culture, though A9 points to evaluator values creating self-fulfilling prophecy. A12 touches on dimensions of organizational culture, remarks on missing elements of metaevaluation.	<p>Overall, analyses as written tend not to address issues of power and relationship; e.g. A3 doesn't challenge sexism.</p> <p>Analyses presume relevance of standard to the evaluation; e.g., the evaluand in A9.2 is not a program. Generalizability of program evaluation standards might be noted or challenged but should not be assumed.</p>	Move the critique beyond methods & procedures to examine underlying values & assumptions that impact accuracy. Comment on challenges and contradictions in adhering to standards. Try to make analyses consistent with Guidelines & Common Errors; e.g., A8.1 only discusses statistical significance; no attention to practical significance. Address societal level issues of racism and sexism as they emerge in case illustrations, impacting both evaluand and evaluation.