

SUMMARY MINUTES ARE DRAFT UNTIL APPROVED AT THE NEXT KTAC MEETING,
SCHEDULED FOR March 28, 2024



Minutes of the January 17-18, 2024 meeting
DRAFT

ATTENDANCE:

KENTUCKY TECHNICAL ADVISORY COMMITTEE (KTAC): Elena Diaz-Bilello (1/18/24 only), Corinne Huggins-Manley, Pete Goldschmidt, Suzanne Lane, Phoebe Winter

KTAC MEMBERS ABSENT: None

KENTUCKY DEPARTMENT OF EDUCATION (KDE): Kevin Hill, Helen Jones, Jenni Larkins, Kevin O’Hair, Michael Prater, Ben Riley, Rhonda Sims, Jennifer Stafford, John Wickizer

KDE GUESTS:

EdMeasure: Bill Auty

Center for Assessment: Brian Gong, Laura Pinsonnault

Caveon: Walt Drane, Marc Weinstein, Joe Kamell

KDE: Karen Dodd

Pearson: Likun Hou, Andrew Thompson, Ed Wolfe, Scott Wilson, Llana Williams, Ashley Hayes

SUMMARY:

The Kentucky Technical Advisory Committee (KTAC) provided advice to the Kentucky Department of Education regarding accountability and reporting options; considerations regarding options for supporting the security of remote testing; received an introduction to the Kentucky United We Learn (KUWL) Council and its work to support development of innovative assessments and accountability systems; advised on procedures to post-equate the Kentucky assessments and methods to evaluate pre-equating; provided advice on human and AI scoring of constructed responses; and suggested ways to strengthen the documentation of the 2023 accountability performance standard setting. The KTAC also set future meeting dates.

Agenda Item: Welcome and Introductions

Presenter: Kevin O’Hair, Academic Program Manager, Office of Assessment and Accountability, KDE

Summary of Discussion: O’Hair started the meeting at 9:00am (Eastern). O’Hair welcomed the KTAC members and presenters. He noted that Rhonda Sims, Associate Commissioner, was attending meetings with the General Assembly, which is in session. He reviewed the background of the establishment of the Kentucky Technical Advisory Committee (KTAC) and reviewed the procedures for complying with Kentucky’s open meetings requirements.

Agenda Item: Review Agenda and Approve Minutes from June 13, 2023 KTAC Meeting

Presenter: Brian Gong, Center for Assessment

Summary of Discussion: Gong invited the KTAC members to introduce themselves and noted that KTAC member Elena Diaz-Bilello was not present, but would join the second day. Gong presented the main topics of the draft agenda; KTAC members had no suggestions to modify the agenda. Gong reminded KTAC members they had reviewed the draft minutes previously. Pete Goldschmidt moved to approve the minutes. Phoebe Winter seconded the motion. There was no discussion to amend the minutes. The minutes were approved unanimously by the KTAC members.

Feedback/Action: The minutes of the June 13, 2023, KTAC meeting were approved by KTAC.

Agenda Item: Kentucky Department of Education Update

Presenters: Kevin O’Hair, Jennifer Strafford, and Kevin Hill, KDE

Summary of Discussion: Jennifer Stafford provided an update of staffing changes in KDE, including the appointment of Interim Commissioner Robin Kinney and an update on the search for a permanent Commissioner. Kevin Hill provided a summary of the 2023 assessment and accountability results portrayed in the media briefing packet. KTAC members had several questions and comments about how to interpret and communicate the results, especially the results of the new accountability system. Kevin O’Hair provided a summary of upcoming test development plans, including shortening the tests and keeping the tests aligned to revised content standards. KTAC members had comments and suggestions about the implications of shortening the tests.

Feedback/Action: KTAC members recommended KDE consider conducting or commissioning analyses of the impact of the new accountability system by various school and student group characteristics with the aim of understanding barriers and identifying successful schools to help other schools. KTAC suggested OAA work with other KDE offices to understand the needs and supports being provided to address priorities, e.g., specific gaps or consistently low or flat achievement. KTAC suggested that a future KTAC meeting

address how KDE and contractors are addressing the desires to shorten state tests and maintain alignment and comparability over time.

Agenda Item: Remote/Virtual Testing and Security

Presenters: Walt Drane, Marc Weinstein, and Joe Kamell, Caveon

Summary of Discussion: Drane, Weinstein, and Kamell introduced themselves and their backgrounds regarding remote/virtual testing and security. Security is an essential aspect of supporting valid interpretations of scores, since violations of security would introduce construct-irrelevant variance. Security also helps maintain program quality and resources, e.g., safeguards being able to develop a planned amount of test items. The Caveon presenters summarized the recent history of remote testing among states, discussed security and fairness concerns, and suggested strategies to help ensure security, fairness, accessibility, and equity. In response to KTAC questions, the presenters indicated that remote testing for states became an approved mode of testing during and following the COVID pandemic when many students were being instructed remotely. There has been an increase in students continuing to be instructed remotely through online learning, and many of these students and their schools would like them to be assessed remotely. KDE mentioned that there is some interest in the state, and this discussion is part of a fact-finding effort on the part of the department.

Feedback/Action: KTAC agreed with Caveon’s advice that states should be very thoughtful and deliberate in any actions regarding remote testing to ensure there is adequate and fair security. KTAC recommended KDE be specific about the purposes and audiences that would be served by remote testing in Kentucky, and make sure security policies and procedures are tailored to those circumstances. KTAC emphasized learning from other states and organizations, including both those that have adopted remote testing and those that have not or have ceased allowing remote testing. KTAC also encouraged KDE to investigate what evidence has been gathered regarding comparability of scores produced under in-person and remote testing conditions.

Agenda Item: Kentucky United We Learn (KUWL) Council Updates

Presenter: Karen Dodd, KDE

Summary of Discussion: Dodd presented the history and activities of the Kentucky United We Learn (KUWL) Council, standing committees, and work groups and how that has evolved over the first year of the CGSA (Competitive Grant for State Assessment) grant. In response to KTAC questions, some examples were given of innovative assessment and accountability approaches being investigated under the KUWL Council: Portrait of a Learner and a “menu” approach wherein KDE would set general parameters and local districts/schools could choose from options within those parameters.

Feedback/Action: The KTAC suggested that the KUWL Council work would benefit from developing a theory of action or logic model, in addition to the “moonshot” and other

goal statements, and the design principles. KTAC also suggested that attention be paid to matching the assessments to the vision of the “vibrant learning experiences,” e.g., how to reflect the “cultural wealth” brought by students in assessments. This is both an area with considerable experience and research (e.g., performance assessments), and an area of active research and development in the measurement field (e.g., culturally relevant assessments). KTAC also suggested that KDE seek out the “lessons learned” from other projects that advocated a similar “system of system” or localization structure, particularly what types of supports were needed to make the implementation successful.

Agenda Item: 2023 Data Results – Post-Equating/Future Pre-Equating Plans

Presenter: Likun Hou, Pearson

Summary of Discussion: Hou presented an overview of the Spring 2023 administration of the Kentucky State Assessments (KSA), focused on a psychometric analysis summary, which included an evaluation of item calibration, description of post-equating procedures and evaluation, and summary of student performance. Hou then presented two options for evaluating pre-equating results, since KDE would like to move to pre-equating to speed up the delivery of assessment result reports to schools and students.

Feedback/Action: KTAC requested more detailed technical documentation in the future (e.g., number, position, item type, common/matrix, and content coding of equating, scored, and field test items by form) to support KTAC’s thorough understanding of the details of the post-equating. KTAC recommended that Pearson include content review of all items flagged (e.g., for drift), and that no items be dropped solely on the basis of not meeting statistical/psychometric criteria. KTAC recommended triangulating potential item drift by using multiple methods, e.g., Robust-Z and displacement. KTAC also recommended that the “purified” anchor set, before being finalized, be checked and documented regarding content alignment to the overall test blueprint.

Given KDE’s desire to move to pre-equating at some point, KTAC reviewed two options for a study of the feasibility of pre-equating; this study could take place initially in the spring of 2024. KTAC members favored a hybrid of the two options and recommended that KDE and Pearson focus both on item-level data (Option 1) and student results (Option 2) in their analyses to inform the decision about whether to proceed with pre-equating. KTAC members favored information based on item stability, and an examination of shifts in proficiency. KTAC members noted that shifts (e.g., differences in classifications of scores at each proficiency level) should be evaluated in terms of consistency, *not* in terms of whether the results produced a higher percentage of students proficient, for example. In fact, KTAC felt strongly that there should be safeguards against deciding based on “favorable results,” such as conducting the evaluation blind. KTAC encouraged checking for consistency between the methods over multiple years, if possible, both before a decision is made to switch to pre-equating and as a check after. If samples of students are used, KTAC advised that much care be taken to ensure the adequate representativeness of the sample; multiple samples could be used as one way to safeguard against biased or inadequate sampling. KTAC noted that post-administration checks of pre-equating results are common after pre-equating has been

adopted as the operational equating method, but is the results are used only to update item calibrations for the item bank.

Agenda Item: Constructed Response Scoring (Evidence and Options)

Presenter: Andrew Thompson and Ed Wolfe, Pearson

Summary of Discussion: Thompson summarized the procedures used by Pearson to help ensure the quality of human scoring of students' written responses to constructed response questions. Wolfe summarized the procedures used by Pearson to help ensure the quality of computer (A.I.) scoring of constructed responses. Both responded to questions and suggestions from KTAC members.

Feedback/Action: KTAC recommended that procedures be included to check for consistency of scoring across years, when items or rubrics are repeated. KTAC suggested that the technical documentation include information alluded to in the presentations, including: how and why target QC criteria were established (e.g., 60% IRR for Writing trait scoring), the nature and number of validity papers (e.g., how representative the pool of validity papers is to what scorers are expected to score), score point distributions for items/traits, cross-distributions of scorers' scores, analyses of scoring quality stability over time/number of responses scored between trainings, percentages of items sent from AI to human scoring (and analyses of why), and why the agreement rates are generally higher (but not always) for scorers to validity papers than between scorers. Comprehensive documentation would include some specific examples of rubrics, student responses, training/qualifying packs (perhaps of released items), etc.

KTAC recommended additional documentation of procedures to ensure adequate representativeness of the student samples used for training AI scoring, and for retraining on the fly. An evaluation of scoring accuracy for targeted groups (e.g., English learners) should also be part of procedures. Could criteria be set for when stratified and/or oversampling is called for? KTAC suggested that the QC criteria be examined and justified with an eye to what the public might find credible to accept AI scoring. KTAC recommended the technical documentation separately document regular training and retraining on the fly, with how Pearson determines the adequate quality of each, particularly if there is incremental acceptable of elements of retraining on the fly.

KTAC encouraged KDE and Pearson to consider future work, including expanding AI scoring to other content areas with constructed responses not currently AI-scored; reducing the amount of human scoring in Writing; using scoring data to refine rubrics; and using knowledge gained from scoring to inform test design and development (e.g., construction of task/item types/templates, general rubrics).

Agenda Item: 2023 Accountability Standard Setting Recap

Presenter: Brian Gong, Center for Assessment

Summary of Discussion: Gong summarized the procedures and outcomes of the 2023 accountability performance standard setting, and referred to the draft report and

communications slide presentation. KTAC complemented KDE for the strengths of the process and documentation, especially the broad representation and transparency in how cut score decisions were made. The discussion highlighted how important communication of the accountability system is; in response to KTAC questions, KDE mentioned several other resources, including pre-release training, simple guides to accountability scores and calculations, “what if” spreadsheet tool, etc.

Feedback/Action: KTAC suggested that the report include a summary of important demographic characteristics of the standard setting panel (e.g., percent female/male; percent from lower/middle/higher performing school systems, etc.). KTAC recommended that versions of the standard setting information be made available to help provide a foundation for the credibility of the system with different audiences, in combination with other resources provided by KDE. Future evaluations might include a question to capture individual panelists’ views (why, what might be better) when they indicate less than strong support for what was recommended by the panel.

Agenda Item: Future KTAC Meeting Dates

Presenter: Brian Gong, Center for Assessment

Summary of Discussion: KDE and KTAC members agreed on three dates and ways that KTAC meetings will be held in the remainder of 2024.

Feedback/Action: The agreed-upon KTAC meeting dates for 2024 are:

- March 28, 2024 (Virtual meeting; tentatively scheduled for 11:00am-5:00pm Eastern)
- June 12-13, 2024 (In-person meeting; location and specific times to be determined)
- October 8, 2024 (Virtual meeting; tentatively scheduled for 11:00am-5:00pm Eastern)