COLORADO
Department of Education

COVID-19 Policy Implications Stakeholder Group

September 30, 2020
Meeting 3

Big Picture: Sequencing Our Work and Today's Agenda

Overview: Bill Language - HB20-1418

(I) **Review the impact of the COVID-19 pandemic and the resulting disruption of the 2019-20 school year**, including student transition to remote learning and the cancellation of the state assessments, accountability, accreditation, and educator evaluation systems for the 2019-20 school year;

(ii) **Discuss how the cancellation of state assessments will impact accountability, accreditation, and educator evaluations during the 2020-21 school yea**r and whether future modifications are needed regarding the accountability, accreditation, and educator evaluation systems as a result of, and in response to, the COVID- 19 pandemic and possible further disruptions; and

(iii) **Make recommendations** regarding whether and how to proceed with state assessments, accountability, accreditation, and educator evaluations during the 2020-21 school year and how the systems can continue to effectively measure student achievement and growth and provide an accurate, credible, and comparable assessment of the quality of the public education system throughout the state following the COVID-19 pandemic.

Overview of Topics

Meeting	**Topics**
Meeting 1 (Aug. 26)	Intros; Background; Context
Meeting 2 (Sept. 16)	Guiding Questions; COVID-19 Impact, Assessments Intro, Educator Evaluations (2020-21)
Meeting 3 (Sept. 30)	Assessments
Meeting 4 (Oct. 14)	Accountability & Accreditation; Educator Evaluations (2021-22)
Meeting 5 (Oct. 29)	Finalize Recommendations

Where Are We Headed?

1. Discussion

2. Identify Emerging Proposal

3. Identify Any Unsatisfied Concerns

4. Collaboratively Modify the Proposal

5. Assess the Degree of Support

6. Finalize the Decision OR Circle Back to Step 1 or 3

Preliminary Consensus Recommendations

?
?
?

Final Product Determined by the Stakeholder Group

Meeting Objectives

1. Discuss Public Participation/Feedback Opportunities

2. Discuss Assessments: Purpose and Technical Overview of Large Scale Testing

3. Discuss and Draft Recommendations for State Content Assessments

Agenda

1:00 - Gather and Welcome from Rhonda Haniford, Associate Commissioner of School Quality and Support

1:05 - Review Agenda and Intro Activity

1:20 - Discussion: Opportunities for Public Participation

1:30 - Presentation: Overview of Technical Requirements of Large Scale Assessments

2:10 - Discussion: Validity Challenges Due to COVID-19 and State Assessments and Purpose, Uses and Unique Qualities

2:25 - Break

2:35 - Presentation: Purpose and Uses of State Content Assessments

3:00 - Discussion and Draft Recommendation(s): CO PSAT/CoAlt 9/10

3:15 -Presentation: Spring 2021 Scenarios: Overview and Process

3:30 - Break

3:40 - Discussion and Draft Recommendation(s): SAT and CMAS

5:00 - Adjourn

Norms

Four Agreements of Courageous Conversations (inspired by Glenn Singleton)

- **Stay engaged**: Staying engaged means "remaining morally, emotionally, intellectually, and socially involved in the dialogue".
- **Experience discomfort**: Discomfort is inevitable. Participants make a commitment to bring issues into the open.
- **Speak your truth**: This means being open about thoughts and feeling and not just saying what you think others want to hear.
- **Expect and accept non-closure** during the process: This agreement asks participants to "hang out in uncertainty" and not rush to quick solutions.

Additional

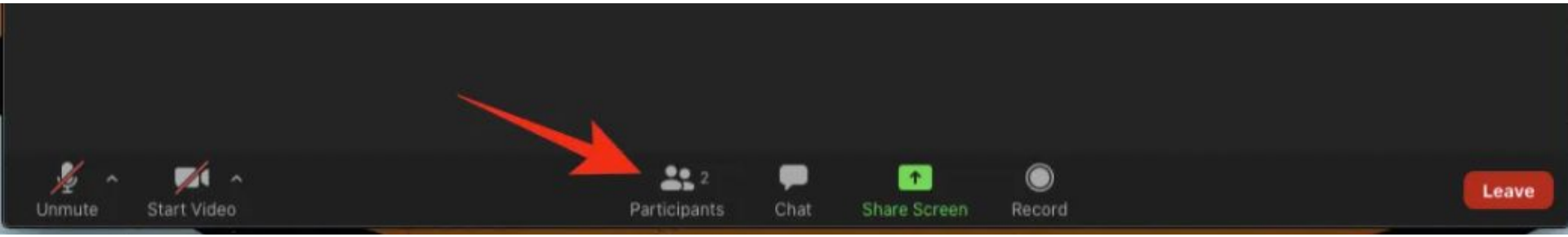
- Mute your microphone, but not yourself.
- Practice patience for the unique conditions under which you're participating.

Meeting Update

Logistics

- Meetings will be recorded and available on the Stakeholder Group webpage

- Zoom Features
  - Raise hand feature
  - Chat box feature



Stakeholders

- Johan Van Nieuwenhuizen, Weld County School District RE-1
- Elizabeth Casillas, Denver Metro Region
- Da'Lisa Hatcher, Third Future Schools - Coperni 2 Charter School
- Peter Hilts, District 49
- Michelle Murphy, Colorado Rural Alliance
- Amy Pitlik, Stand for Children
- Kym LeBlanc-Esparza, Archuleta School District 50 JT
- Bret Miles, Colorado Association of School Executives
- Jennifer Holladay, Denver Public Schools
- Luke Ragland, Ready Colorado
- Jason Westfall, Colorado Association of BOCES
- Cheri Wrench, Colorado Association of School Boards
- Christina Ethier, Association for Colorado Education Evaluators and Cherry Creek School District
- Paul Freeman, Roaring Fork School District RE-1
- Chris Gibbons, Colorado League of Charter Schools and STRIVEPrep
- Carol Eaton, Technical Advisory Panel and Jefferson County School District
- Stephanie Perez-Carrillo, Colorado Children's Campaign
- Amie Baca-Oehlert, Colorado Education Association
- Jen Walmer, Democrats for Education Reform

Public Participation

Public Participation Ideas

What are your ideas for public participation in the Stakeholder Group process?

- Public Feedback Form
- Focus Groups for feedback
- Open Comments - written or live
- Others

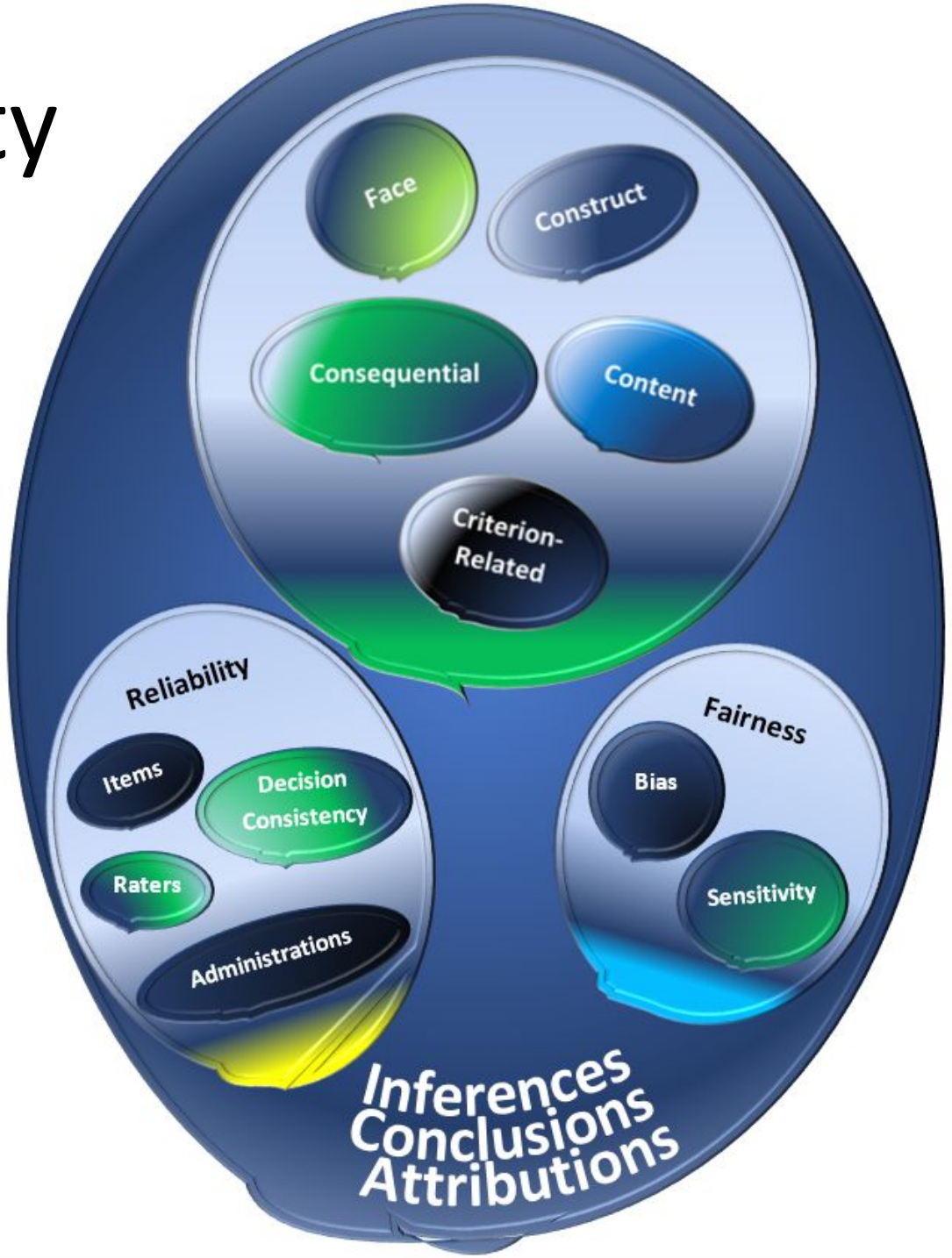
Spring 2021
State Content Assessments

Advance Organizer

- Overview Presentation of Technical Requirements of Large Scale Testing
- Break
- Purpose and Uses of State Content Assessments
- Back to CO PSAT/CoAlt 9/10
- ESSA and Colorado Content Assessment Requirements
- Spring 2021 Scenarios
  - Flowcharts
  - Preliminary recommendations

Overview Presentation of Technical Requirements of Large Scale Assessments

Validity


Face
Construct
Consequential
Content
Criterion-Related
Reliability
Items
Decision Consistency
Raters
Administrations
Fairness
Bias
Sensitivity
Inferences
Conclusions
Attributions

Peer Review Requirements

- **Section 1: Statewide System of Standards and Assessments**
- Section 2: Assessment System Operations
- **Section 3: Technical Quality – Validity**
- Section 4: Technical Quality – Other
- Section 5: Inclusion of All Students
- Section 6: Academic Achievement Standards and Reporting

Technical Requirements of Large Scale Assessments Reliability and Validity

Reliability:

- Not about correctness
- All about consistency/reproducibility
  - Across items
  - Across raters
  - Across administrations (time and place)

Technical Requirements of Large Scale Assessments Reliability and Validity

- Most important aspect of reliability: Decision consistency
  - If the student takes:
    - different forms of the test
    - at different times (no new instruction between tests),
    - rated by different people, and
    - receives the same performance level,

    the assessment has decision consistency.

  - High decision consistency is necessary for substantive actions to take place based on a test.

Technical Requirements of Large Scale Assessments Reliability and Validity

- Most important aspect of reliability: Decision consistency
  - Decision consistency can be higher across students based on aggregated scores than on a single score.
    - Why? Another concept: true score
    - Theoretically, every student has a true score that represents the absolutely accurate level of achievement of the student in that domain.
    - But, every time a student takes a test, there's some noise, some error. It may be caused by a momentary distraction or reduced effort (a tickle in the nose, a breakup with a girlfriend that morning, etc.) that leads to a student getting an item wrong that they do actually know. Or, a student may just guess well (lucky?) on a particularly tough item.
    - With group scores, the error tends to balance out. Some students score a little higher than their true score; some score lower. It averages out, leading to higher decision consistency.

Technical Requirements of Large Scale Assessments Reliability and Validity

Relationship between reliability and validity:

- Reliability is a necessary but not sufficient condition for validity.
- A test can be highly reliable (consistent) but have very low validity. Both reliability and validity are not conditions of absolutes. They exist as a matter of degree.
  - A test could have high reliability, but have low validity because it is scoring consistently inaccurately.
  - A test could have high reliability, but be missing some other key validity characteristics.

Technical Requirements of Large Scale Assessments Reliability and Validity

More aspects of validity (also continuum-based): To what extent,

- Face validity- does the test look like what you would expect for a test claiming to measure X?
- Content validity- are the items on the test fairly representative of the entire domain (knowledge and process, balance of content, cognitive complexity, depth and breadth, cognitive processes, and subdomain structure) the test seeks to measure ?
- Construct validity - does the test measure what it claims or purports to measure?
- Criterion-related - does the test correlate with a target outcome or measure (concurrent or predictive)?
- Consequential validity – does the test lead to positive or negative social/societal (or personal) consequences?

Technical Requirements of Large Scale Assessments
Reliability and Validity
Fairness

Fairness

- Bias
- Sensitivity

Technical Requirements of Large Scale Assessments Reliability and Validity

- What is a key feature that contributes to a valid assessment and resulting scores?
    - Standardization of all processes and procedures from initial test design through reporting.
- Significant issue for this year: adherence and fidelity to standardized administration and security policies and procedures should remote testing be an option.

Administration Options by Assessment

Assessment	In-Person Administration	Remote Administration
PSAT/SAT	X	Not Available
CMAS	X	Being Investigated
CoAlt (DLM)	X	Possibly Available to Home-Bound Students
CoAlt Science and Social Studies	X	Possibly Available to Home-Bound Students
ACCESS for ELLs	X	Possibly Available to Home-Bound Students

Technical Requirements of Large Scale Assessments Reliability and Validity

A test is considered “valid” based on a body of evidence collected across time demonstrating that it possesses the necessary characteristics and the inferences made based on it are reasonable and justifiable.

Peer Review Requirements

- **Section 1: Statewide System of Standards and Assessments**
- Section 2: Assessment System Operations
- **Section 3: Technical Quality – Validity**
- Section 4: Technical Quality – Other
- Section 5: Inclusion of All Students
- Section 6: Academic Achievement Standards and Reporting

Peer Review Requirements

- **Section 1: Statewide System of Standards and Assessments**
  - **1.1 State adoption of academic content standards for all students**
  - **1.2 Challenging academic content standards**
  - **1.3 Required assessments**
  - **1.4 Policies for including all students in assessments**
  - 1.5 Meaningful consultation in the development of challenging state standards and assessments

Peer Review Requirements

- Section 2: Assessment System Operations
  - **2.1 Test design & development**
  - **2.2 Item development**
  - 2.3 Test administration
  - 2.4 Monitoring test administration
  - 2.5 Test security
  - 2.6 Systems for protecting data integrity & privacy

Peer Review Requirements

- **Section 3: Technical Quality – Validity**
  - **3.1 Overall Validity, including validity based on content**
  - **3.2 Validity based on cognitive processes**
  - **3.3 Validity based on internal structure**
  - **3.4 Validity based on relations to other variables**

Peer Review Requirements

- Section 4: Technical Quality – Other
  - **4.1 Reliability**
  - 4.2 Fairness & accessibility
  - **4.3 Full performance continuum**
  - **4.4 Scoring**
  - **4.5 Multiple assessment forms (comparability)**
  - 4.6 Multiple versions of an assessment (comparability)
  - **4.7 Technical analysis & ongoing maintenance**

Peer Review Requirements

- Section 5: Inclusion of All Students
  - 5.1 Procedures for including students with disabilities
  - 5.2 Procedures for including English learners in academic content assessments
  - 5.3 Accommodations
  - 5.4 Monitoring test administration for special populations

Peer Review Requirements

- Section 6: Academic Achievement Standards and Reporting
  - **6.1 State adoption of academic achievement standards for all students**
  - **6.2 Achievement standards setting**
  - **6.3 Challenging & aligned academic achievement standards**
  - 6.4 Reporting

6.4 Reporting
Non-accountability Assessment Results

For any assessment (non-accountability) results that are released, what additional data or information would support their appropriate use?

Contextual information on how students learned:

- Schooling approach during the year?
- Attendance/engagement?
- Availability of technology (device and internet)? How has that changed from last spring?
- Home support/environments? (Where is that line into too sensitive, too personal?)
- Curriculum coverage?

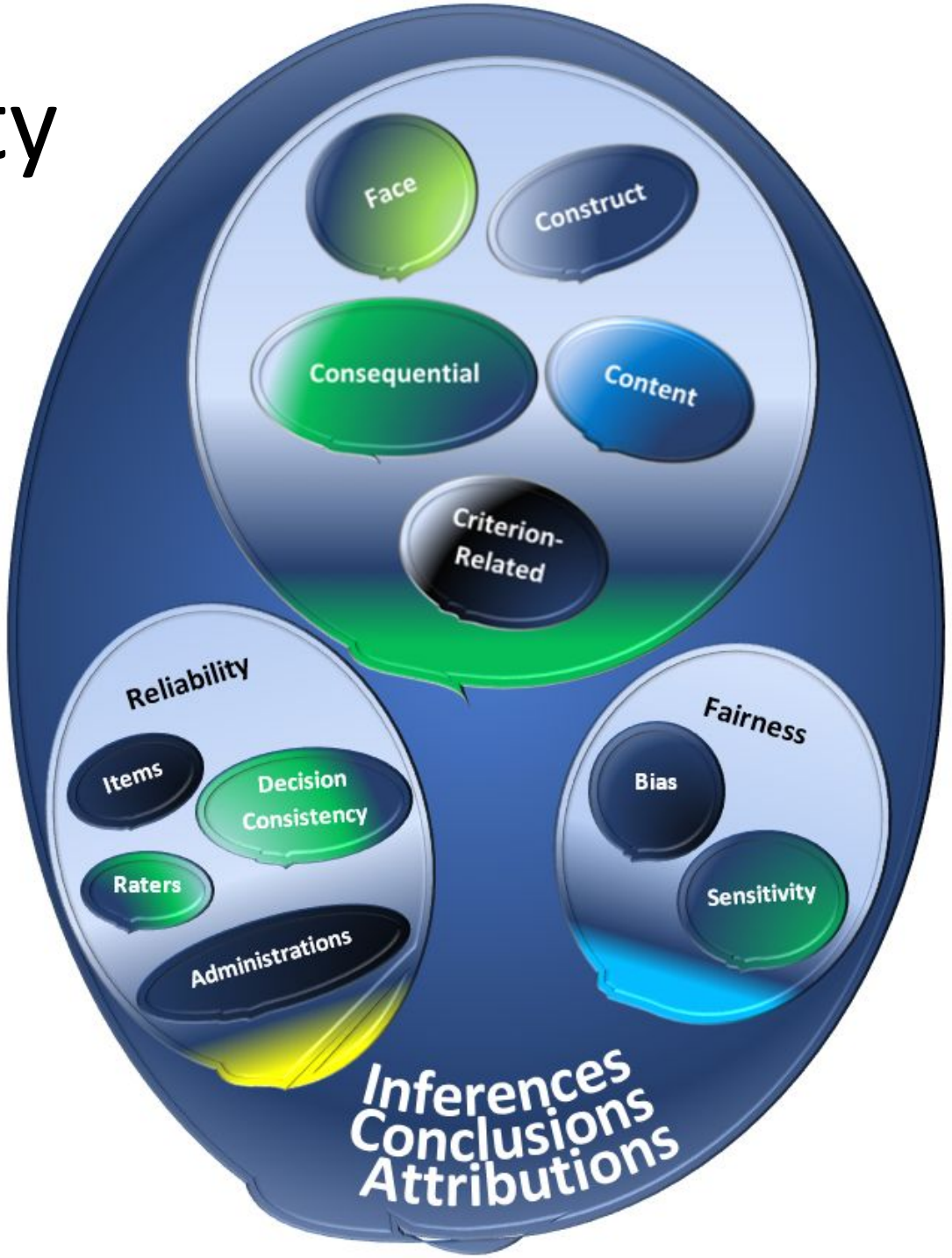
We are expecting that additional interpretation guidance will be necessary.

What can we do to reduce the perception that any released results are sanctions?

When you hear "testing won't be valid," what does that mean?

- We've talked about the test itself and its resulting scores.
- Questions of validity are often concerns with interpretation, conclusions that are drawn, and attributions of causality rather than something inherent to the test itself.
- Margie scored 752 on her CMAS math assessment.
  - Based on the above, how valid are the below statements? Which are well-supported, which require additional evidence, and which would you not recommend making?
    - "Margie performed at Met Expectations on CMAS this year."
    - "Margie scored higher than the average student score in the state."
    - "Margie is less capable than previous years' students."
    - "Margie worked hard this year."
    - "COVID-19 didn't impact Margie."
- Remember the advantage of aggregated scores.

Validity


Face
Construct
Consequential
Content
Criterion-Related
Reliability
Items
Decision Consistency
Raters
Administrations
Fairness
Bias
Sensitivity
Inferences
Conclusions
Attributions

Discussion

Which of the technical requirements pose the biggest challenge to validity due to COVID-19?

Who are the users of state assessment results? How are they used? What unique information do they provide?

BREAK

Purpose and Uses of State Content Assessments

Why do we test?
What's unique about the state (CMAS/CoAlt) assessments?
What unique information does it provide?
What is unique about what people want to know about effectiveness of responses to COVID-19?

Purpose and Uses of State Content Assessments

Purpose of State Content Assessments:

> The state content assessments are intended to be one component of the more comprehensive local assessment systems. They typically serve as **uniform** measures across the state of end-of-year achievement of the **Colorado Academic Standards** (CAS) at the individual student, school, district and state levels; spotlight inequities; and help target getting the right supports for students.

For some districts, they are the only standardized assessments that address higher order and critical thinking skills, including independently composed writing.

Results may be used by a variety of stakeholders to inform different questions and activities.

Purpose and Uses of State Content Assessments

Uses for students, parents, educators, schools, community and policymakers have to be decoupled.

- <u>Student and Parent</u>
  - Results typically:
    - Provide a shared state-wide indicator of individual student mastery of end-of-year standards, including areas of relative strength and weakness
    - Allow comparisons of individual results against the school, district and state
      - An external measure based on grade-level standards that is given to all students across the state can help parents answer the following types of questions:
        - How well is my student/student's school performing in comparison to the neighboring school, another district or the state?
        - Is my student is on track for success for the next grade level?
        - Is my student being sufficiently challenged or supported?
  - Results may be a piece of information used to inform school selection
    - Many parents trust school ratings when looking for places to live and send their students to school. Ratings incorporate state assessment results.

CMAS Preliminary Student Performance Report

CMAS: English Language Arts
Spring 2020

Preliminary Student Report - Confidential

FIRSTNAME003 Y LASTNAME 003 (2019030110)

Test Details

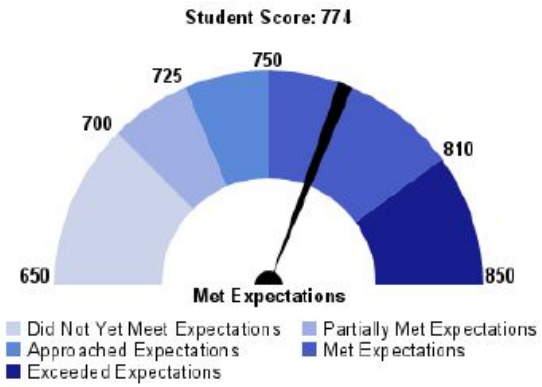
Test Name
Grade 03 ELA

Grade
03

Responsible District
SAMPLE DISTRICT NAME

Responsible School
SAMPLE SCHOOL NAME (CO-5555-4444)

Preliminary - How Did FIRSTNAME003 Perform Overall?



Grade 3 English Language Arts

This preliminary score report provides information about your student's mastery of skills and concepts in the academic standards which are the basis for instruction in Colorado schools. Your student's performance on this test is represented by a scale score and a performance level.

For more information on the CMAS assessment program, visit http://www.cde.state.co.us/assessment/cmas.

Performance Level Descriptor – Met Expectations

FIRSTNAME003 **Met Expectations** and is on track for the next grade level. Students in this level typically demonstrate the following:

Reading
- With very complex text: the ability to be generally accurate when asking and/or answering questions, showing general understanding of the text when referring to explicit details and examples in the text.
- With moderately complex text: the ability to be generally accurate when asking and/or answering questions, showing general understanding of the text when referring to explicit details and examples in the text.
- With readily accessible text: the ability to be mostly accurate when asking and/or answering questions, showing understanding of the text when referring to explicit details and examples in the text.

Writing
Written Expression: students typically address the prompts and provide development of ideas, including when drawing evidence from multiple sources, while in the majority of instances demonstrating purposeful and mostly controlled organization. Students typically:
- Develop the topic and/or narrative elements using reasoning, details, text-based evidence, and/or description.
- Develop topic and/or narrative elements in a manner that is mostly appropriate to the task and purpose.
- Demonstrate purposeful organization that is mostly controlled and may include an introduction and/or conclusion.
- Use linking words and phrases, descriptive words, and/or temporal words to express ideas with clarity.

Knowledge and use of Language and Conventions: students typically demonstrate command of the conventions of Standard English consistent with edited writing. There are errors in grammar and usage that may occasionally impede

CMAS Individual Student Performance Report

Confidential Student Performance Report

Colorado Measures of Academic Success

Student: FIRSTNAME018 LASTNAME018
SASID: 2019040834 Birthdate: 01/09/2006
School: SAMPLE SCHOOL NAME (4444)
District: SAMPLE DISTRICT NAME (5555)

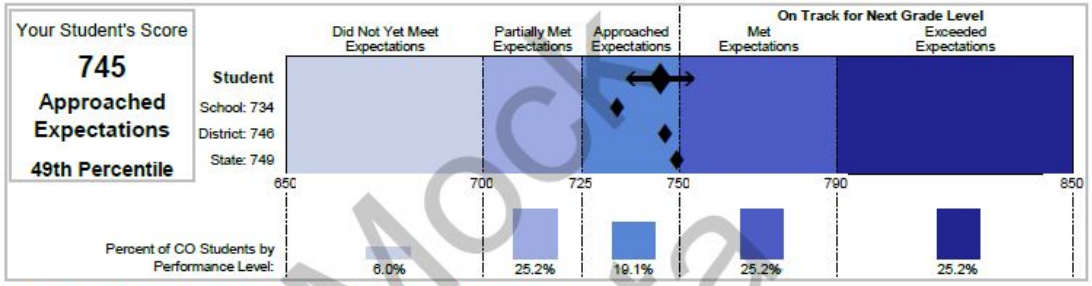
Spring 2019

English Language Arts/Literacy — Grade 4

This score report provides information about your student's performance on the Colorado Measures of Academic Success (CMAS) English Language Arts/Literacy test.

- Your student's performance is represented by a scale score, a performance level, and a percentile rank. Scores are placed on a scale so that student performance can be compared across years.
- On the graph, scale scores are represented by diamonds. The arrows around your student's diamond show the range of scores your student would likely receive if the assessment was taken multiple times.
- School, district, and state information is provided so that you can compare your student's performance to the performance of others. The percentage of students in each performance level across the state is reported below the graph.
- Dotted lines show where the range of scores is divided into performance levels.
- You are encouraged to discuss this report with your student's teacher.

Your Student's Score
745
Approached Expectations
49th Percentile



Mock Data

Performance Level Descriptor - Approached Expectations

Students who **Approached Expectations** may benefit from additional support to meet expectations at the next grade level and they typically demonstrate the following:

In **Reading**, the pattern exhibited by student responses indicates:

- With very complex text: the ability to ask and/or answer questions with minimal accuracy, showing minimal understanding of the text when referring to explicit details and examples in the text.
- With moderately complex text: the ability to be generally accurate when asking and/or answering questions, showing basic understanding of the text when referring to explicit details and examples in the text.
- With readily accessible text: the ability to be mostly accurate when asking and/or answering questions, showing understanding of the text when referring to explicit details and examples in the text and when explaining inferences drawn from the text.

In **Written Expression**, students typically address the prompts and provide basic development of ideas, including when drawing evidence from multiple sources, while in the majority of instances demonstrating organization that sometimes is controlled. Students typically:

- Develop topic and/or narrative elements in a manner that is general in its appropriateness to the task and purpose.
- Demonstrate some organization.
- Include some linking words and phrases, descriptive words, and/or temporal words, limiting the clarity with which ideas are expressed.

In **Knowledge and use of Language Conventions**, students typically demonstrate basic command of the conventions of Standard English consistent with edited writing. There are few patterns of errors in grammar and usage that impede understanding, demonstrating partial control over language.

4

To view the full version of the performance level descriptors (PLDs), visit:
http://www.cde.state.co.us/assessment/grade_4_english_language_arts_plds.

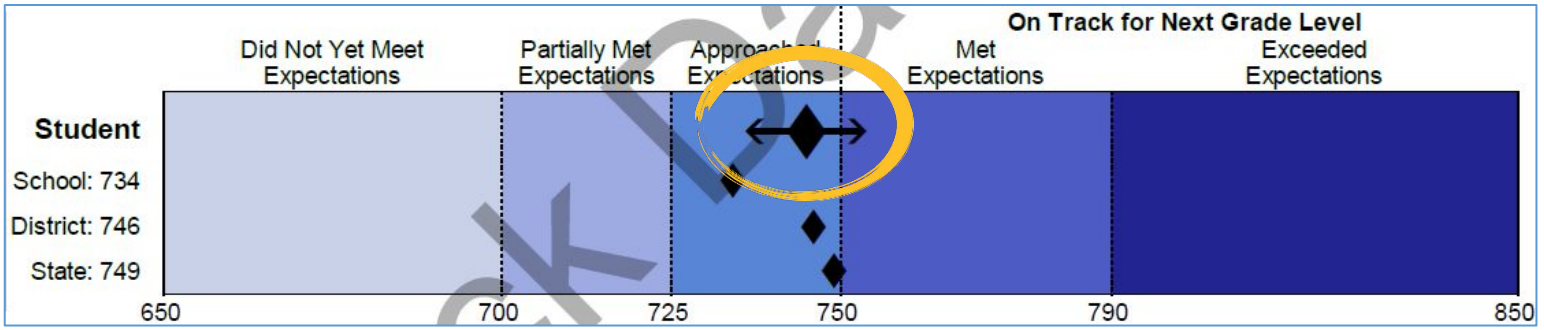
Purpose
This report describes your student's mastery of the Colorado Academic Standards in Reading and Writing.

For more information on the CMAS assessment program, visit
http://www.cde.state.co.us/assessment/cmas

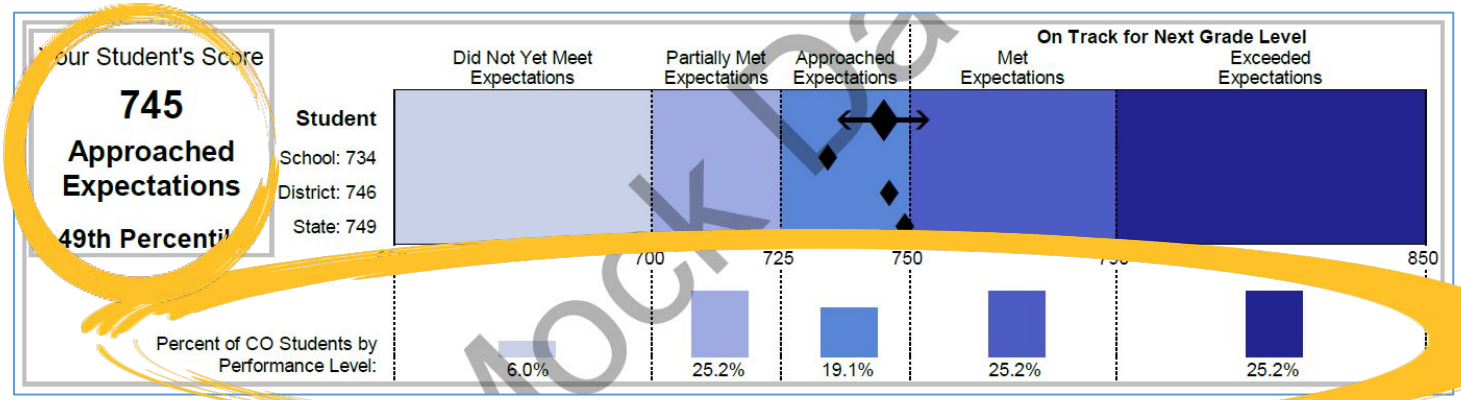
12192018-Z9999999-5555-4444 - 0000000

Math and ELA Overall Scale Score and Performance Level

- This student
  - Performed better than school average, but below the district and state averages



- Percentage of students within the state at each of the performance levels



Mock Data

Purpose and Uses of State Content Assessments

- School and District
  - Results typically provide a piece of data that:
    - May be used as a prompt for further investigation at the student, classroom, school and district levels
    - May support districts/schools in reviewing and developing goals for the performance of their students, including subgroups
      - District, school and classroom trends from the previous year can be used to inform district and school level programming as well as instructional plans for the next school year.
      - Teachers can use results to gain better understandings of incoming students' individual strengths and weaknesses at the beginning of the school year to help provide targeted support.
    - May indicate that a review of programs, curricula, materials and/or scope and sequence may be appropriate
    - May inform the evaluation of instructional approaches

Purpose and Uses of State Content Assessments

- State
  - Aggregated results are a source of data that:
    - Are incorporated into the state accountability system allowing for school and district comparisons of
      - Aggregated student demonstration of their mastery of the CAS
      - Aggregated student growth toward mastery of the CAS
    - Are used as prompts for further investigation at the school, district and state levels
    - Support the state in reviewing and developing its goals
    - Provide information to inform and evaluate legislative and policy decisions
    - Are used in determining schools to be recognized and identified for a variety of reasons, including serving as an example for other districts/schools
    - Serve as a signal of values and priorities
- Other uses of results:
  - Research purposes
  - Development of reports
  - Informing community and organization efforts

Purpose and Uses of State Content Assessments

What is unique about the state assessments within the standardized components of local assessment systems?

- Content and scores aligned specifically to Colorado standards (criterion-referenced), especially CMAS (all content areas) and CoAlt (Science and Social Studies)
- CSLA as non-transadapted accommodated form of CMAS ELA
- ELA, including CSLA, passages reviewed by Colorado educators
- Written and reviewed by Colorado educators (CMAS (all content areas) and CoAlt (Science and Social Studies))
- Scoring points on constructed response items recommended by Colorado educators
- Cut scores for performance levels as indicators of mastery of CAS grade level expectations recommended by Colorado educators
- Number and percent of constructed response items/points is significantly higher than on most off-the-shelf interim assessments
- School, district and state comparative data provided

Purpose and Uses of State Content Assessments

Consider potential increased interest given COVID-19

- What questions do people have about achievement and growth in Colorado now and in two years?
  - Will they want to estimate the impact of how districts responded to COVID-19 on student learning?
  - Will they want to estimate how effectively different standards were addressed at different grade levels?
  - Will they want to estimate differential impact on subgroups?
- What is the "unit" of interest (students, classroom, school, district, state, etc.)?
- What data will we need to have to answer those questions?
- What will they want to do with that information?

- What are the state's intended uses, priorities and capacity?

Back to Colorado Assessments: CO PSAT/CoAlt 9/10

Colorado State Content Assessments (C.R.S. § 22-7-1006.3)

Colorado Measures of Academic Success (CMAS)	Colorado Alternate Assessments (CoAlt)	CO PSAT/SAT SAT – grade 11* **PSAT – Grades 9 and 10****
Math*	Math (DLM)*	Math
ELA*	ELA (DLM)*	Evidenced-based Reading and Writing
Science*	Science*	**Writing (SAT optional essay)*** **
Social Studies**	**Social Studies****	

* Required by Colorado law and federal law
** Required by Colorado law
*** Allowed by Colorado law

Consensus Process

1. Discussion
Q&A: WHOLE GROUP
DISCUSSION: Whole Group

2. Identify Emerging Proposal
Whole Group

3. Identify Any Unsatisfied Concerns
Whole Group

4. Collaboratively Modify the Proposal
Whole Group

5. Assess the Degree of Support
Whole Group

6. Finalize the Decision OR Circle Back to Step 1 or 3
Whole Group

CO PSAT/CoAlt 9/10

CO PSAT 9/10

- Administered in 9th and 10th grades
- Sets expectations and provides experience for the redesigned SAT
- Connected to free review resources
- Used in state and federal accountability (achievement and/or growth)
- ~300 schools with 37,0000+ eleventh graders registered to take PSAT/NMSQT this fall (Wed., Oct. 14th and Thurs., Oct. 29th)
- College Board has modified some of its standardized administration and security procedures in response to COVID-19. These changes may inform spring administration options.

Issues that are Solely within State Purview

PSAT/CoAlt 9/10:

- What are the advantages and disadvantages of administering CO PSAT/CoAlt this spring?
- Given COVID-19, assuming hybrid or at least some in-person instruction and/or assessment is taking place to allow for school building-based administration, should CO PSAT/CoAlt be administered this spring?

1. Discussion

2. Identify Emerging Proposal

3. Identify Any Unsatisfied Concerns

4. Collaboratively Modify the Proposal

5. Assess the Degree of Support

6. Finalize the Decision OR Circle Back to Step 1 or 3

ESSA and Colorado Content Assessment Requirements

ESSA and Colorado Assessment Requirements

The content assessments must:

- Be the **same academic assessments** used to measure the achievement of **all public school students**
- Be administered to all public school students (except for students with the most significant cognitive disabilities participating in alternate assessment) in:
  - Math and English language arts in grades 3-8 and once in high school (9-12)
  - Science once each in elementary, middle and high school (10-12)
  - All other content area assessments administered at the discretion of the state
- **Be aligned to the full breadth and depth of the state** (college and career ready) **standards (CRITERION-REFERENCED)**
  - Assess **higher-order thinking skills and understanding**
  - **Provide indication about whether students are performing at grade level (at least three performance levels)**
- Be of high quality, valid, reliable and consistent with relevant, nationally recognized professional and technical testing standards as evaluated and determined through **Peer Review**
- **Provide appropriate accommodations** for students with disabilities and English learners

ESSA Assessment Requirements - Results

Assessment results must:

- Include reports that are provided to parents, teachers, and school leaders in an understandable and **uniform** format, in a language that parents can understand
- Enable results to be **disaggregated** at the **State, LEA**, and **school level** by: major racial and ethnic groups, economically disadvantaged as compared to the not economically disadvantaged, children with disabilities as compared to children without disabilities, English proficiency status, gender, and migrant status (unless N size is too small)

Break

Spring 2021 Scenarios

Spring 2021 State Content Assessments: COVID-19 Effects

"If at all possible, states should plan to administer statewide summative tests in the spring of 2021 on schedule. A clear picture of statewide performance will encourage state policymakers and appropriators to help schools through what is likely to be a very difficult period. States need the information to drive planning and resource allocations. Comparable data, disaggregated by race, income, geographical region and other student subgroups, is more important than ever given the disparate impact of school closures on students. And parents and the public deserve to know where student learning stands.

"We should mount an effort at a census survey of what kids know and are able to do at some point over the next academic year because there's nothing like hard data to ring an alarm bell that's louder than the one that's currently ringing," says Ho of Harvard. "Actual evidence of how much kids have declined, potentially, on average, and how much gaps have grown from previous years needs to be understood and not just guessed at." Such data could drive additional, badly needed investments in education."

Olson, J. (July 2020). *Blueprint for Testing: How Schools Should Assess Students during the COVID Crisis*

Federal Department of Education's Current Position On Spring 2021 Testing

"Accountability aside, we need to know where students are so we can address their needs…Our instinct would not be to give those waivers…There are so many benefits to testing and it allows for some transparency about how schools are performing and the issues we need to address, that our instinct would be to decline those waivers."

Jim Blew, Asst. Secretary of Education (July 24, 2020)

"[I]t is now our expectation that states will, in the interest of students, administer summative assessments during the 2020-2021 school year, consistent with the requirements of the law and following the guidance of local health officials. As a result, you should not anticipate such waivers being granted again."

Betsy DeVos, Secretary of Education, in Letter to Chief School Officers (September 3, 2020)

Secretary's Letter to Chief State School Officers on Assessment for Spring 2021

So if and until new directives are delivered…

The Department is working with its vendors and Colorado districts to prepare for at least some testing in spring 2021.

The feasibility of remote administration to meet the standardization expectations of typical high stakes state-wide administrations has yet to be demonstrated large scale.

With some significant disclaimers, individual student scores resulting from remote administrations that may be cautiously interpreted locally among student, parent and educators may be possible and desirable.

Students Learning Remotely
Options for Testing

- Seeking lessons learned by districts both last spring and this fall with their own attempts at remote administration with their district assessments.
- Hearing that at least one district has indicated that even under mostly remote instruction, it intends to bring their students in safely to test for both district and state testing.
- Hoping to learn from this fall's PSAT/NMSQT and SAT school day administrations.
- Needing a better understanding of district approaches for their students with disabilities and English learners, even when the majority of students may be receiving their instruction remotely.

Students Learning Remotely
Issues that Need to be Addressed

- Comparability: Will scores resulting from remote administrations be comparable to each other and to those from standard school site administrations?
  - The likelihood of being able to develop technical processes to aggregate scores from remote administrations with each other or with those resulting from non-remote administrations is proving more challenging.
- Equity: Will all students have access to devices and reliable internet for remote administrations?
- Equity: Will all students with disabilities have access to the accommodations remotely that they would if they were tested at a school site?
- Equity: Will students be able to create a low disturbance testing environment?
- Security: Will there be adequate safeguards in place to ensure that work completed represents independent effort on the part of the student?
  - We might be able to have a system that would lock a student out if they try to access support on the testing device, but many students have multiple devices. Visual monitoring is probably not an option for Colorado due to privacy concerns.
- Security: Devices used for remote testing may not have the typical app installed. Browser-based testing may be a possibility. How do you provide technical support to remotely testing students?

Reliable Enough: What conditions have to be met to answer this year's questions and this year's intended uses?

If we encourage the use of any provided assessment results from remote administrations to focus more on monitoring and support, how much can we relax some of the standardization of the remote administrations to be able to provide "reliable enough" information on individual student mastery of the standards and areas warranting closer attention for particular student groups?

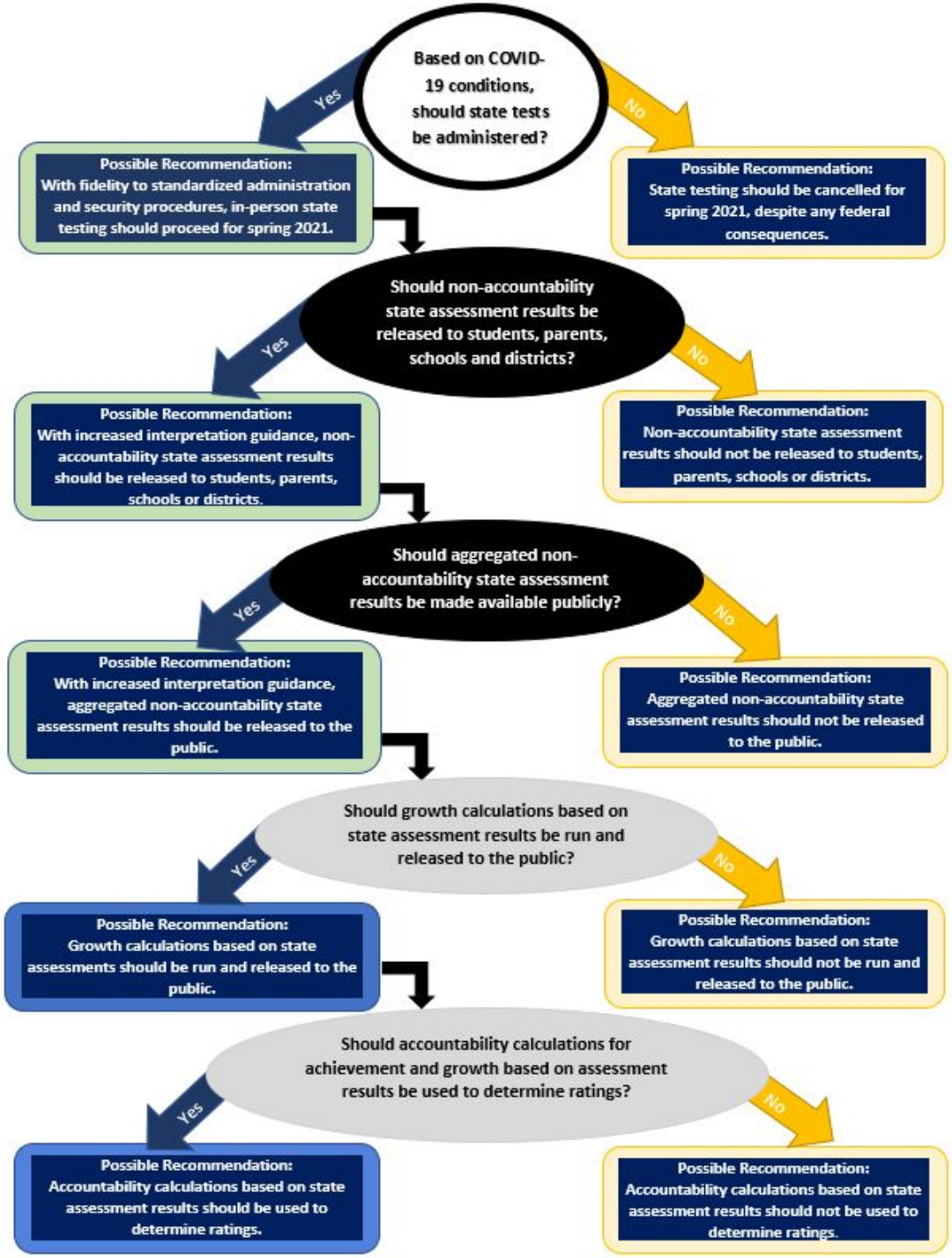
Potential Spring 2021 COVID-19 Conditions

- Condition A: Fully/Vast Majority In-Person Learning
- Condition B: Blended or Hybrid Learning
- Condition C: Fully Remote Learning

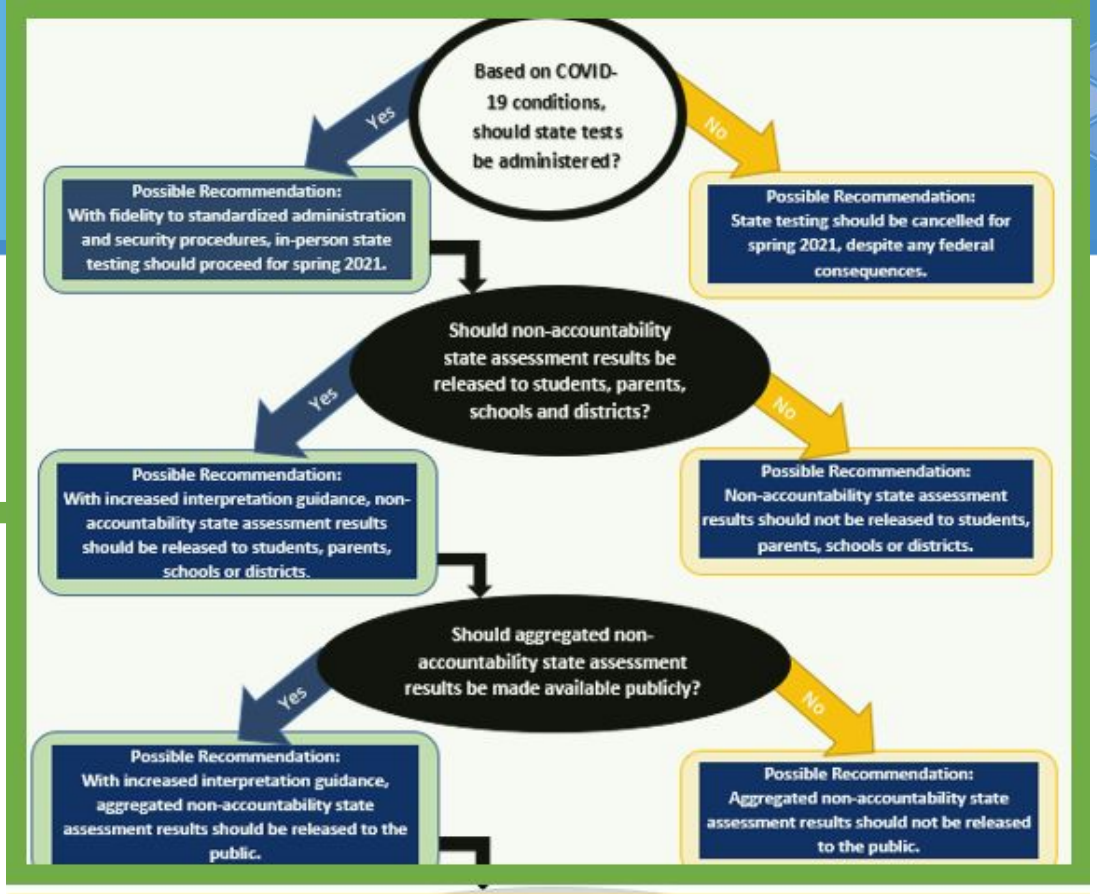
Consider Conditions A-C under current federal expectations*

*Based on current communications from the federal department of education, federal waivers from testing are not expected to be available at this time. In addition to considering Conditions A through C under current federal expectations, the conditions may also be considered under alternate expectations in which federal waivers from testing are available in particular circumstances.

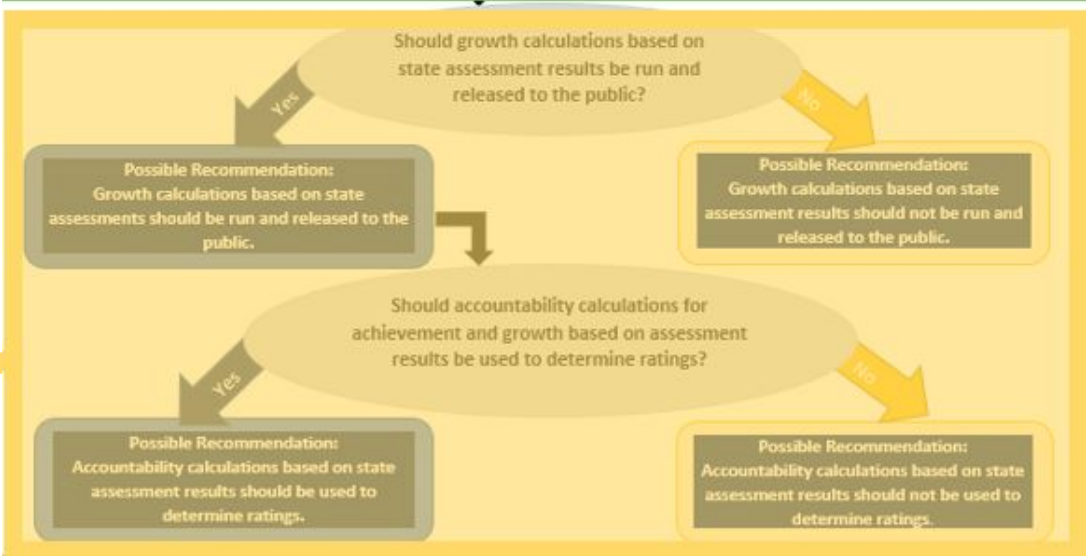
Potential Spring 2021 COVID-19 Conditions Flowchart



Assessment Considerations



Accountability Considerations

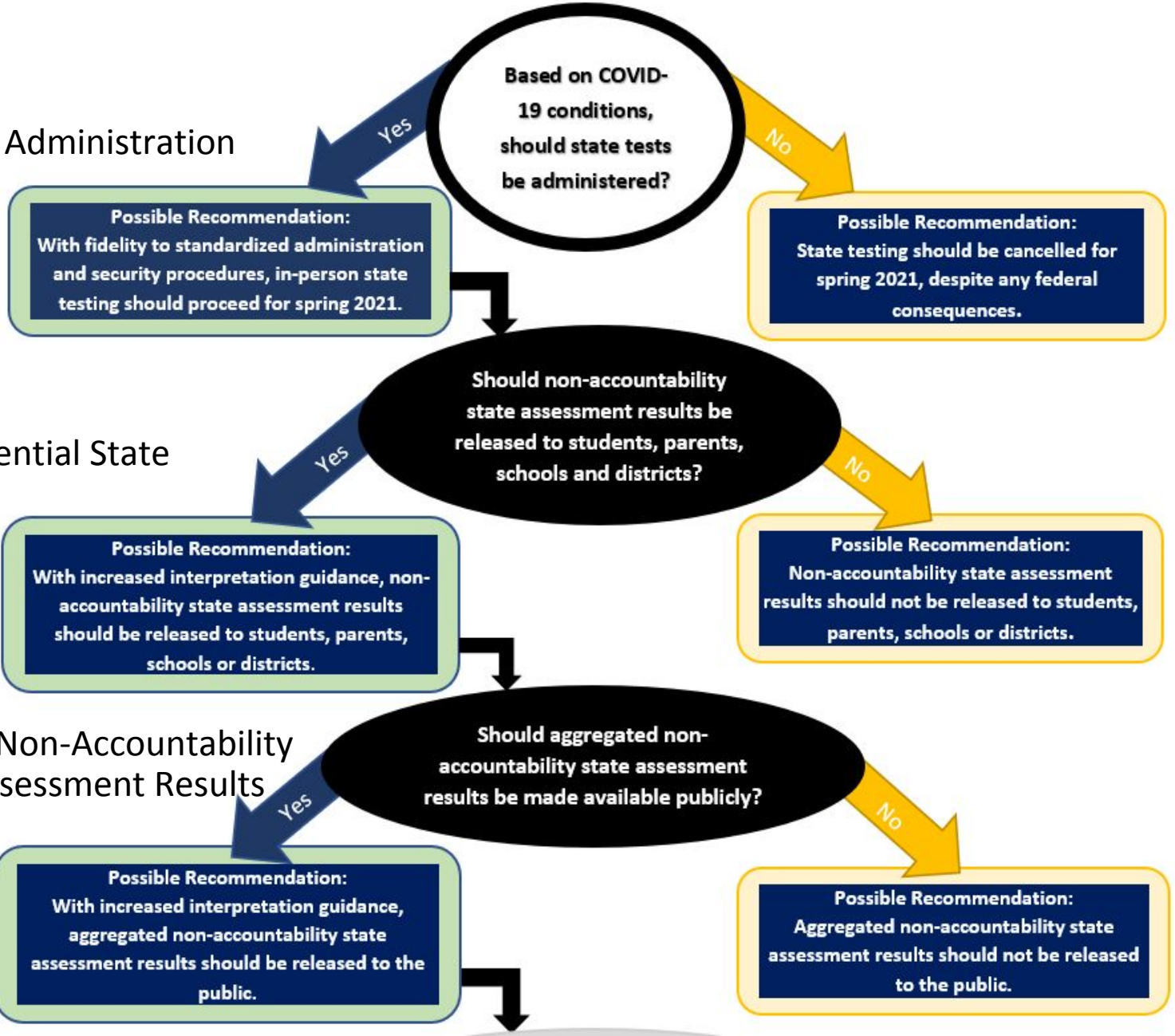


Assessment

1. State Assessment Administration

2. Release of Confidential State Assessment Results

3. Public Release of Non-Accountability Aggregated State Assessment Results



Guiding Assessment Questions
Full/Vast Majority In-Person Learning

- For Spring 2021, if the vast majority of students have returned to in-person learning, should **SAT/CoAlt** be administered following local and state health guidelines to students who are receiving in-person instruction and/or assessment?
  - If yes, should non-accountability state assessment results be released to students, parents, schools and districts (for educational planning purposes)?
    - If yes, should aggregated non-accountability state assessment results be made available publicly?

- For Spring 2021, if the vast majority of students have returned to in-person learning, should **CMAS/CoAlt** be administered following local and state health guidelines to students who are receiving in-person instruction and/or assessment?
  - If yes, should non-accountability state assessment results be released to students, parents, schools and districts (for educational planning purposes)?
    - If yes, should aggregated non-accountability state assessment results be made available publicly?

Guiding Assessment Questions
Blended or Hybrid Learning

- For Spring 2021, if students and schools are engaged in blended or hybrid learning, should **SAT/CoAlt** be administered following local and state health guidelines to students who are receiving some in-person instruction and/or assessment?
    - If yes, should non-accountability state assessment results be released to students, parents, schools and districts (for educational planning purposes)?
        - If yes, should aggregated non-accountability state assessment results be made available publicly?

- For Spring 2021, if students and schools are engaged in blended or hybrid learning, should **CMAS/CoAlt** be administered following local and state health guidelines to students who are receiving some in-person instruction and/or assessment?
    - If yes, should non-accountability state assessment results be released to students, parents, schools and districts (for educational planning purposes)?
        - If yes, should aggregated non-accountability state assessment results be made available publicly?

Guiding Assessment Questions
Fully Remote Learning

- For Spring 2021, if feasible, should **CMAS** be administered to students who are receiving <u>all of their instruction and assessment remotely</u> if reasonably technically sound scores can be produced for the purpose of local interpretation and decisions? (Scores more than likely would not be aggregated with scores from in-person administrations.
  - If yes, should non-accountability state assessment results be released to students, parents, schools and districts (for educational planning purposes)?
    - If yes, but comparability requirements for aggregation are not met, should aggregated non-accountability state assessment results from in-person administrations be made available publicly separate from remote administration results?

Spring 2021 State Assessment Administration and Non-Accountability Results Discussion

In consideration of potential spring 2021 COVID-19 conditions …

1. Based on COVID-19 conditions, should state tests be administered?
2. Should non-accountability state assessment results be released to students, parents, schools and districts?
3. Should aggregated non-accountability state assessment results be made publicly available?

1. Discussion

2. Identify Emerging Proposal

3. Identify Any Unsatisfied Concerns

4. Collaboratively Modify the Proposal

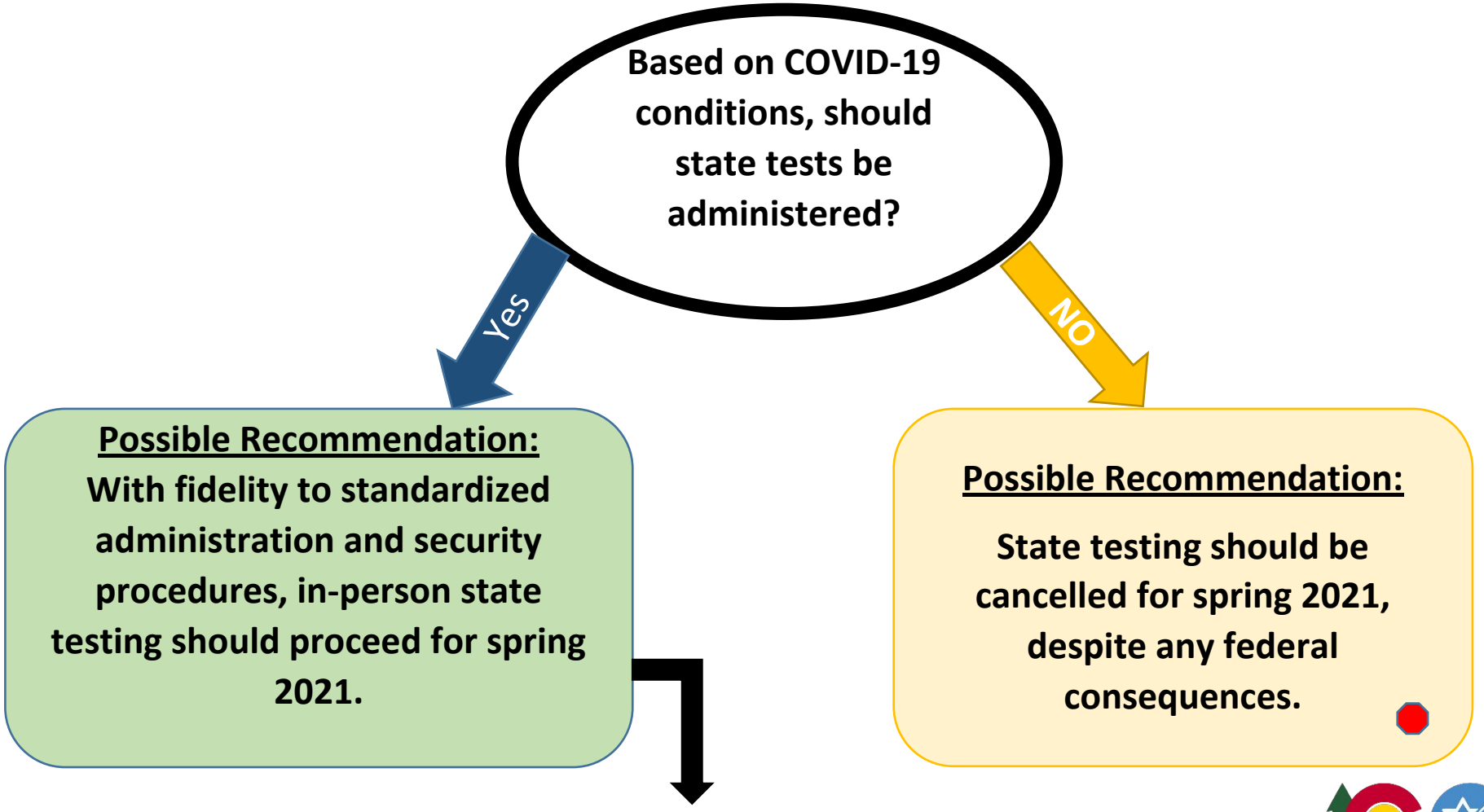
5. Assess the Degree of Support

6. Finalize the Decision OR Circle Back to Step 1 or 3

Condition A: Fully/Vast Majority In-Person Learning

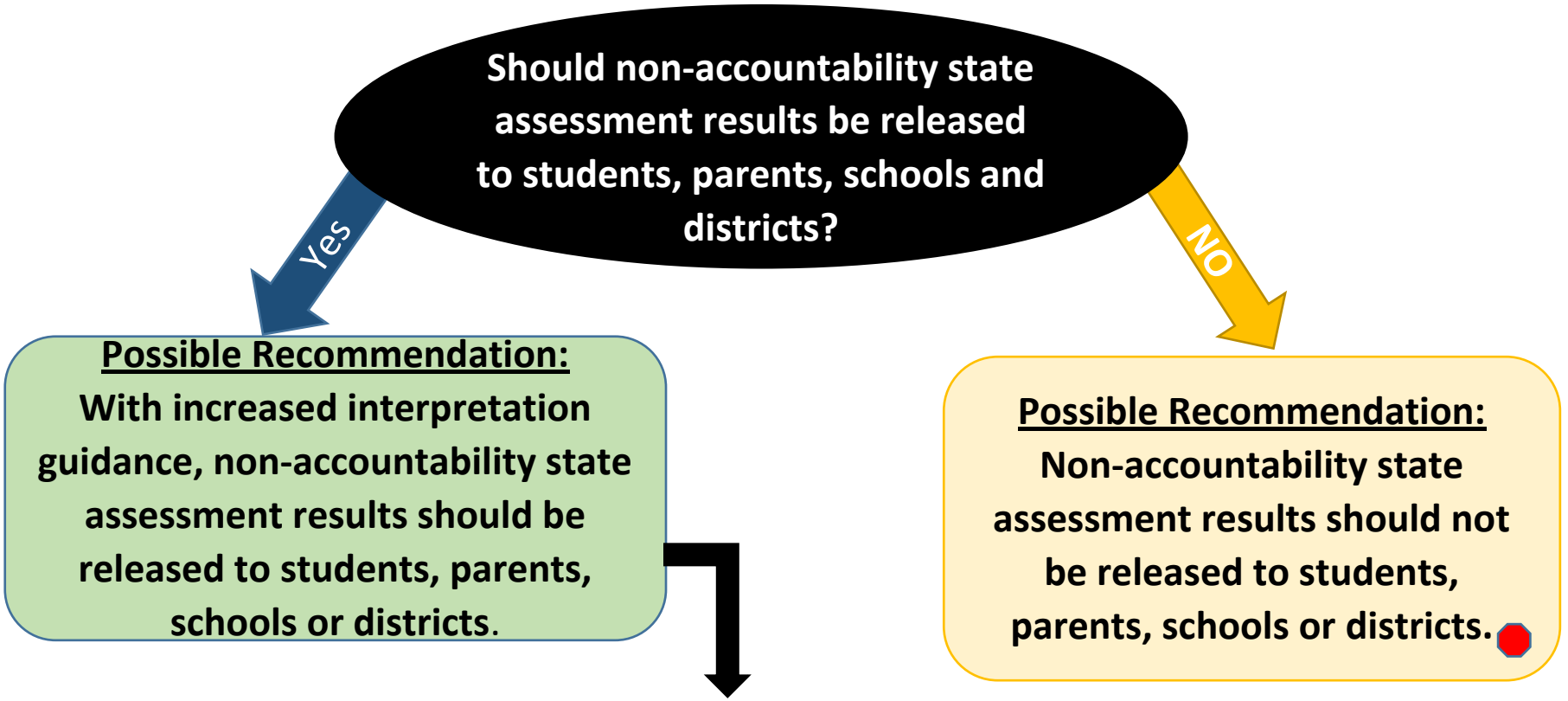
- COVID-19 is controlled and the vast majority of students have returned to in-person learning by the beginning of the testing window.
- Content, constructs and scoring of individual student results remain (reliable and) valid to answer the question: “Has the student demonstrated mastery of the grade level standards?”
- Fidelity to standardized test administration and security procedures can be prepared, trained and monitored for when assessments are administered to students at in-person settings.

Condition A: Fully/Vast Majority In-Person Learning



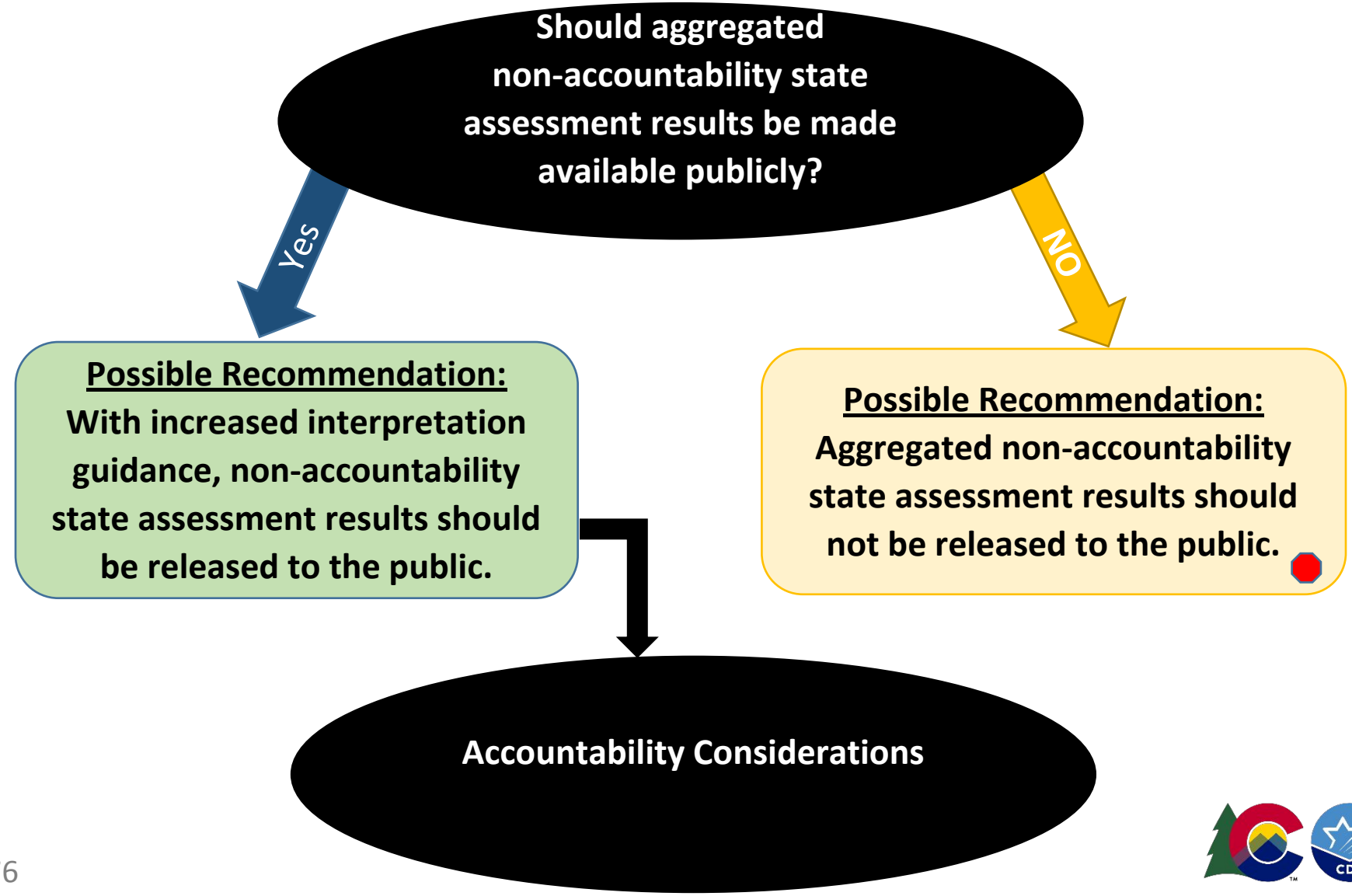
Condition A: Fully/Vast Majority In-Person Learning

Release of Spring 2021 State Assessment Results to students, parents, schools and districts



Condition A: Fully/Vast Majority In-Person Learning

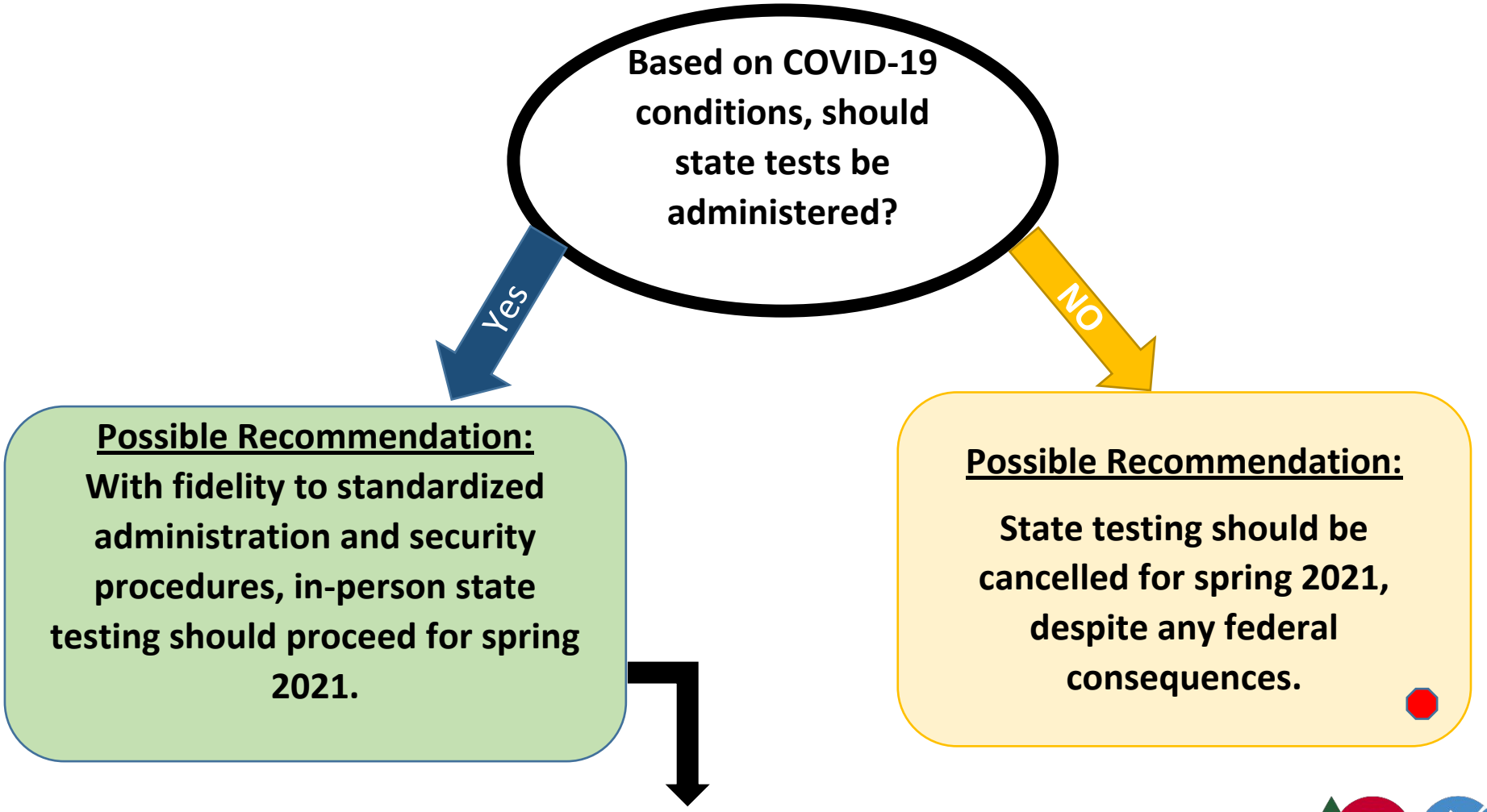
Release of Spring 2021 State Assessment Results to the Public



Condition B: Blended or Hybrid Learning

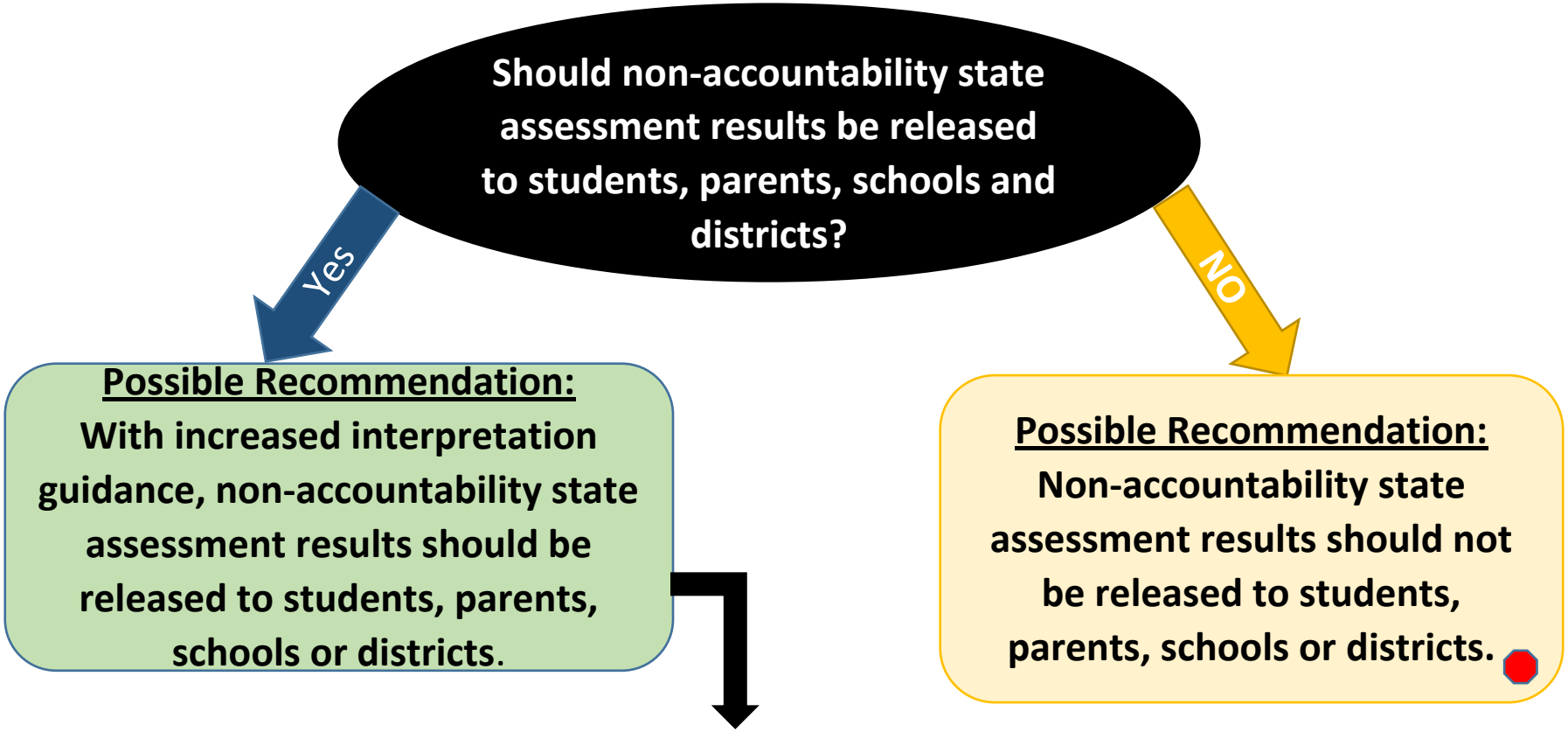
- Schools are in a blended or hybrid model at the beginning of the testing window. Some students may be attending classes in-person, some may be attending classes remotely, and some may be attending a combination of in-person and remote classes.
- Content, constructs and scoring of individual student results remain (reliable and) valid to answer the question: “Has the student demonstrated mastery of the grade level standards?”
- Fidelity to standardized test administration and security procedures can be prepared, trained and monitored for when assessments are administered to students at in-person settings.

Condition B: Blended or Hybrid Learning



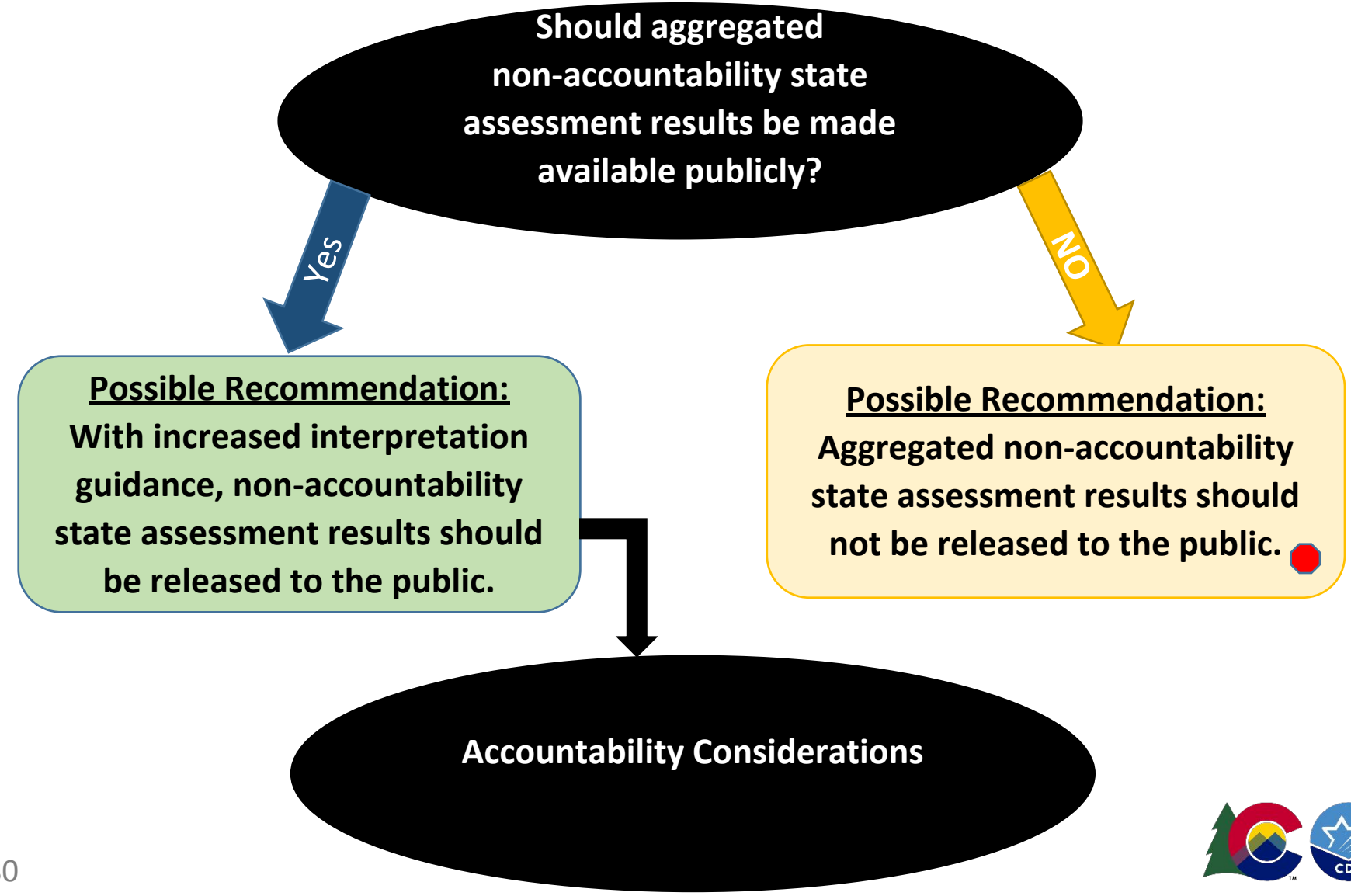
Condition B: Blended or Hybrid Learning

Release of Spring 2021 State Assessment Results to students, parents, schools and districts



Condition B: Blended or Hybrid Learning

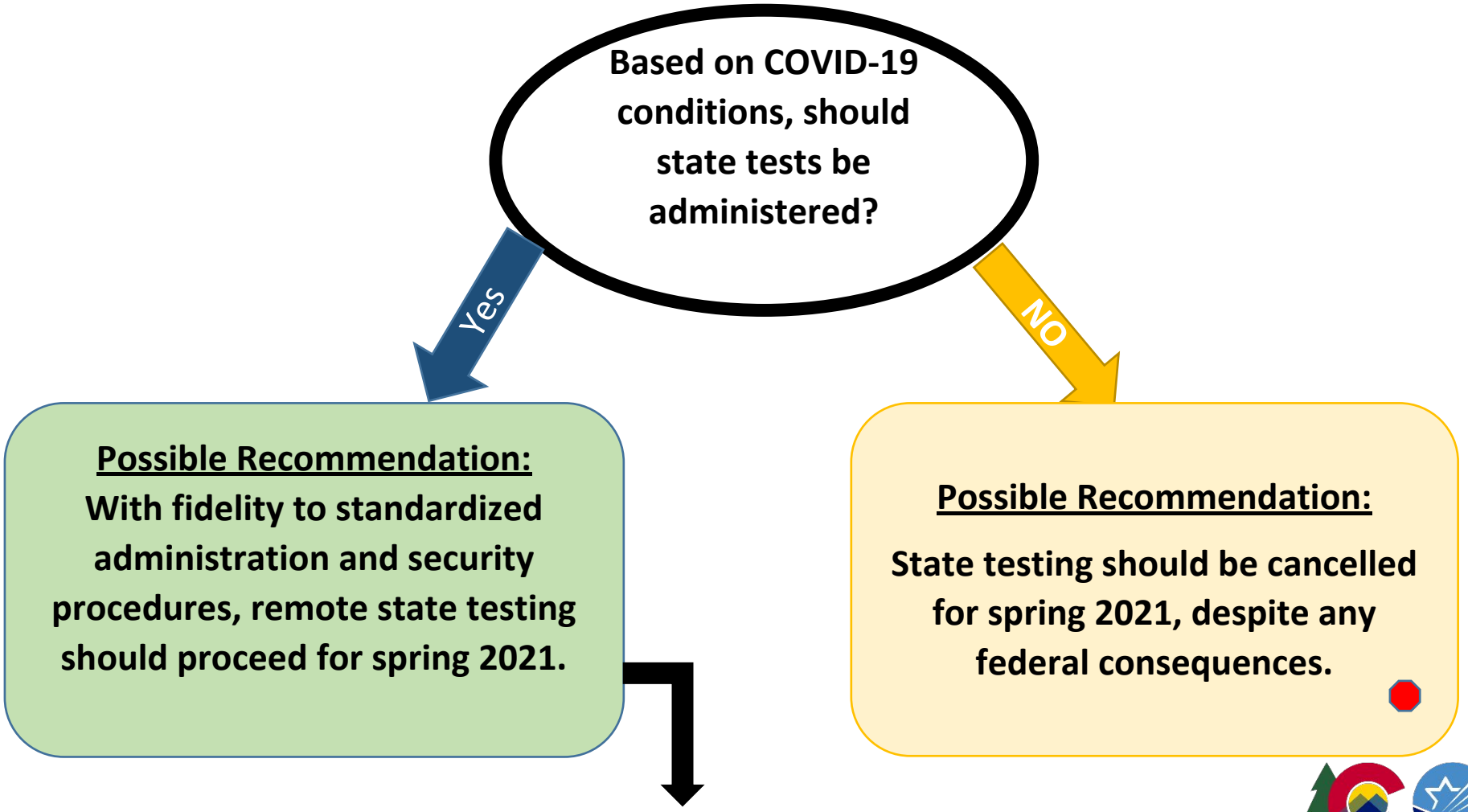
Release of Spring 2021 State Assessment Results to the Public



Condition C: Fully Remote Learning

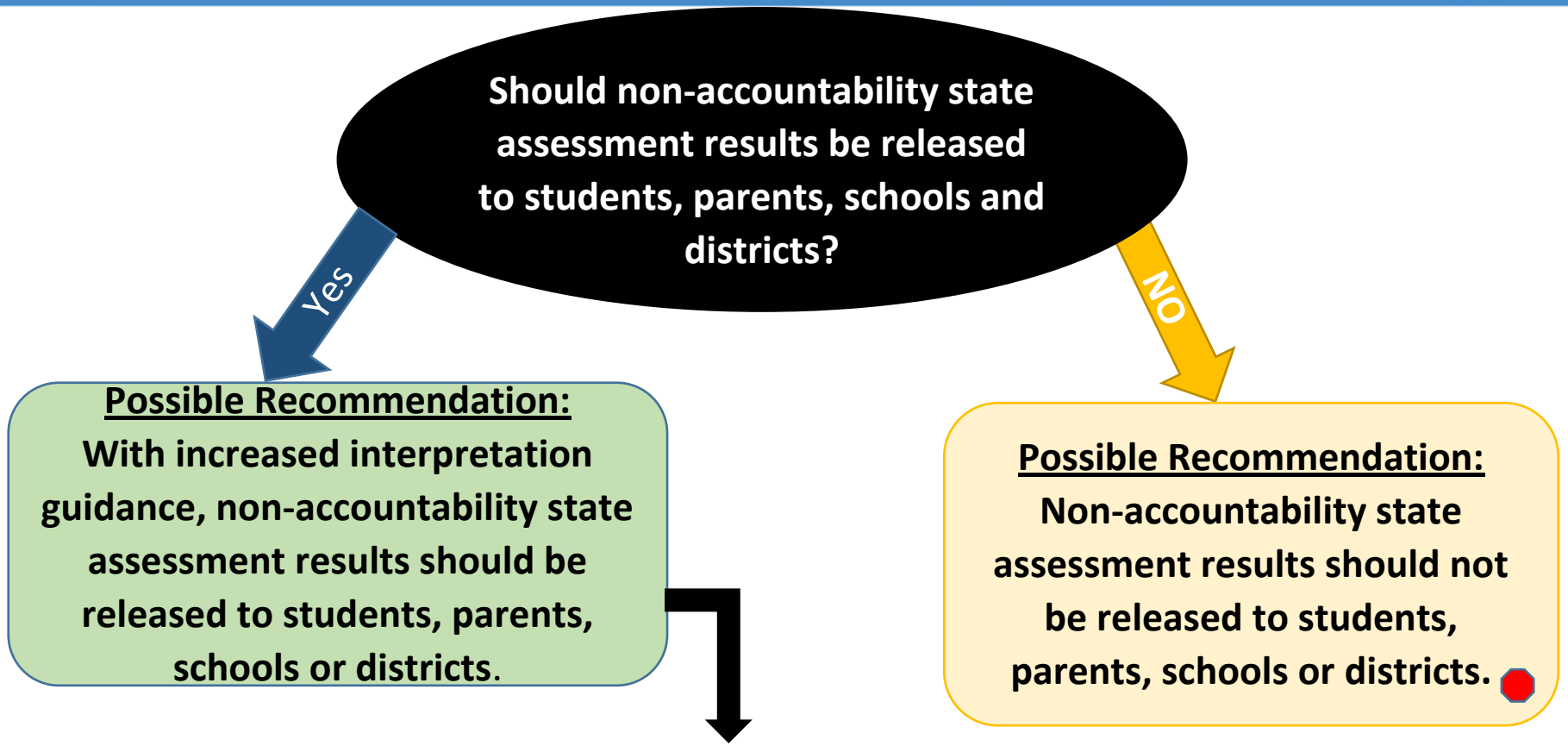
- COVID-19 is not controlled at the beginning of the testing window and the vast majority of students are learning remotely, similar to spring 2020 learning.
- Content, constructs and scoring of individual student results remain (reliable and) valid to answer the question: “Has the student demonstrated mastery of the grade level standards?”
- Fidelity to standardized test administration and security procedures may not be adhered to, prepared or monitored for in remote settings.
- Individual student level results may be produced that allow for meaningful use among student, parent and educator. At this time, there is not a clear answer for how to aggregate scores from these types of administrations.

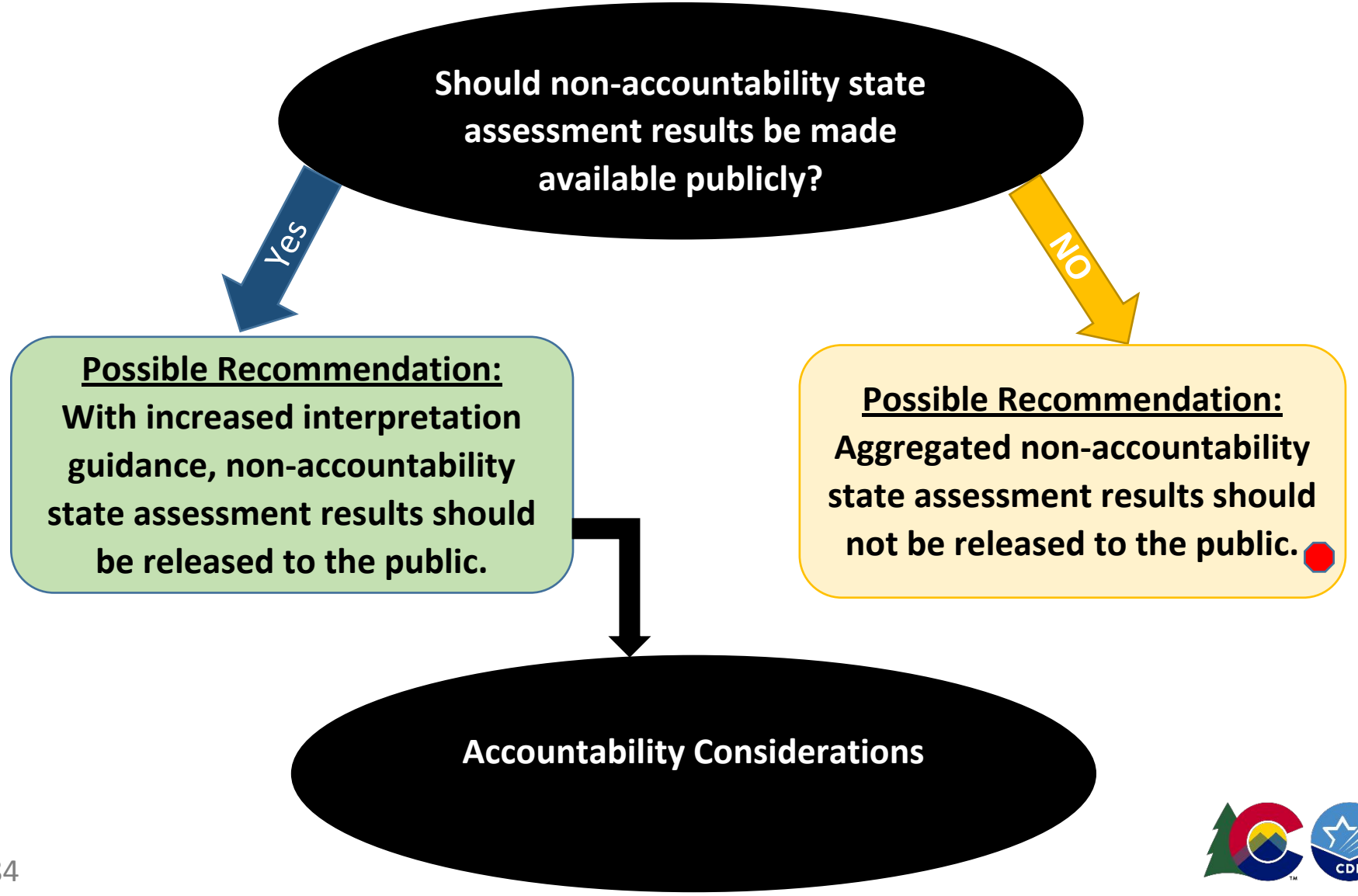
Condition C: Fully Remote Learning



Condition C: Fully Remote Learning

Release of Spring 2021 State Assessment Results to students, parents, schools and districts



Condition C: Fully Remote Learning

Release of Spring 2021 State Assessment Results to the Public