



One Size Doesn't Fit All: Dealing with Exceptionality in State Accountability Systems

Erika Landl, Carla Evans, and Chris Domaleski

Center for Assessment

*Prepared for the Council of Chief State School Officers CCSSO
Accountability Systems and Reporting (ASR) State Collaborative*

June 2020





Table of Contents

Section One	2
Introduction and Context	2
Purposes of School Accountability Systems.....	2
What is Exceptionality?	3
The Solution Space	5
Organization.....	7
Section Two	7
Designing Alternative School Accountability Systems: What, Why, and How?	7
Characteristics of Alternative Schools: The What.....	7
Rationale for Alternative School Accountability Systems: The Why	11
The Design of Alternative School Accountability Systems: The How	12
Section Three	21
Exceptional Circumstances in Accountability	21
Implications of Exceptional Circumstances.....	21
Strategies for Addressing Exceptionality.....	23
Final Thoughts	25
References	26



Section One

Introduction and Context

The writer James Thurber is credited with the saying, “There is no exception to the rule except that every rule has an exception.” We think there is a lesson in this adage that applies to contemporary school accountability systems.

Contemporary school accountability initiatives coordinated by state education agencies are principally, but not exclusively, influenced by the parameters of the federal Every Student Succeeds Act (ESSA). ESSA requires that states annually hold schools accountable for school performance based on academic achievement, English language proficiency, graduation rate and other factors. These systems must produce information that is reliable, valid, and comparable across schools. Consequently, states have produced standardized systems that rely on uniform indicators and rules. But what happens when one size doesn’t fit all? As we will discuss, there are many factors related to school characteristics or context that may impede a state’s ability to produce or implement a standard system for all schools. These circumstances threaten not only the state’s ESSA system, but other accountability initiatives designed to produce information intended to be regarded as comparable within or across years.

In this paper, we address the process and rationale for addressing exceptions in otherwise standardized school accountability systems. While our paper is strongly influenced by systems developed in response to the current reauthorization of the Elementary and Secondary Education Act—the Every Student Succeeds Act (ESSA)—we do not constrain our remarks to current federal law. Indeed, we intend for the issues and ideas explored in this paper to apply broadly to most systematic school accountability initiatives, especially those implemented at the state level.

Purposes of School Accountability Systems

We begin by reflecting on the purposes of school accountability systems and the extent to which those purposes are supported or thwarted when applied to exceptional schools and circumstances. There are at least four primary purposes of school accountability systems (Supovitz, 2009). These are: motivate system change, provide data necessary for decision-making, align system components, and signal system accountability to stakeholders.

Motivate system change is based on the idea that accountability can motivate improvement by attaching either positive or negative incentives to results. This positions accountability as an action lever and assumes that schools have the capacity and means to improve, but need the incentive to do so. Many would argue this position is based on a deficit model because it assumes educators lack the motivation to improve and incentives alone will address this shortfall.

Provide data useful for decision-making positions the information supplied by accountability system indicators as an essential ingredient for school improvement decision-making. This assumes that indicators are (a) able to measure system progress because they are sensitive enough to detect changes (either positive or negative) in the system, (b) at an appropriate level of granularity or specificity, and (c) useful for guiding school improvement efforts and evaluating whether school improvement efforts are working.



Align system components means that school accountability efforts can play a major role in spurring the alignment of major components of the educational system. In other words, system-wide improvement efforts can be supported best if efforts and initiatives such as the alignment of standards, curriculum, instruction, and assessment reinforce each other.

Signal system accountability to stakeholders is based on the ideas of public transparency and answerability. School accountability systems signal important values to stakeholders such as academic achievement, academic growth, closing of achievement gaps, and school attendance.

What is Exceptionality?

In order for an accountability system to support any of the aforementioned purposes states must identify and address potential areas of exceptionality. What do we mean by exceptionality? Broadly, it refers to characteristics of schools or atypical circumstances that could impede efforts to collect accurate, intended information about the performance of one or more schools. These factors are elaborated below and summarized in Figure 1.

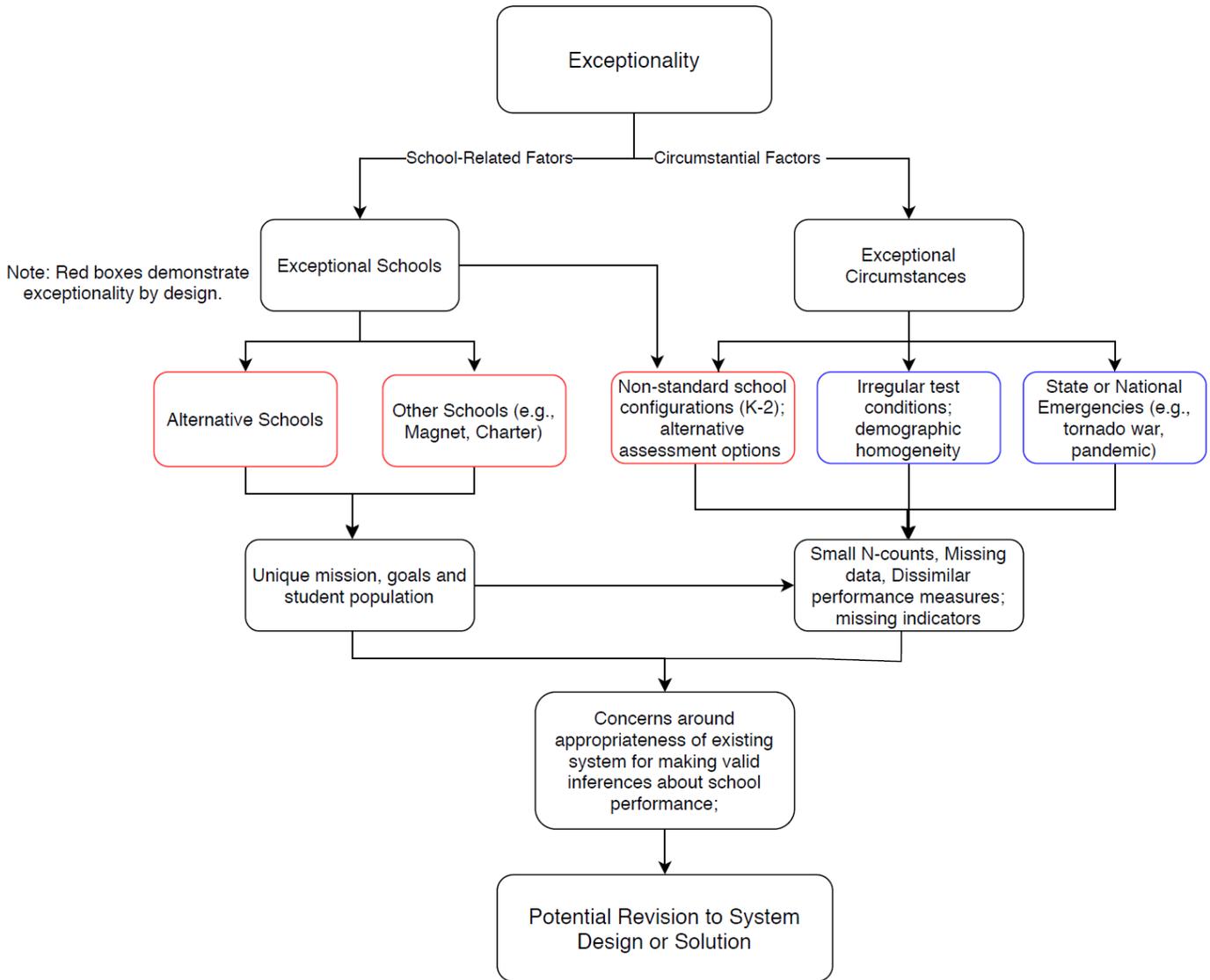
Exceptional schools are *schools* with distinct student populations, missions or characteristics. There are many types of exceptional schools including schools for the deaf and blind, some magnet and charter schools, and those commonly referred to as alternative schools. We will unpack the term “alternative school” in more detail in section two, but we use it as shorthand to describe the broad range of schools that focus primarily on serving students who have not been successful in a traditional school environment. In our experience, the opportunities and challenges associated with designing and implementing school accountability systems for alternative schools are among the most prominent and pressing for states.

Exceptional circumstances are *conditions* that impede or complicate the application of a standardized school accountability model to one or more schools. While exceptional schools often demonstrate exceptional circumstances the classification is not limited to these schools. A common exceptional circumstance is when a school has an insufficient N-count to support the calculation of one or more indicators within the accountability system. For example, a school having a small number of English learners may not be able to calculate the English language proficiency indicator required under ESSA. Another example is when a system is missing assessment data because students opted not to test or there was a testing irregularity due to cheating. While these are common situations they represent “exceptional circumstances”, as defined, because they do not allow the system to be implemented as intended and/or call into question the appropriateness of the model for some schools.



Figure 1

Potential Pathways from Exceptionality to System Re-Design



We acknowledge that classifying exceptions as pertaining to schools and circumstances is not mutually exclusive. For example, a K-2 school may be thought of as an “exceptional school” rather than dealing with an “exceptional circumstance” because it supports a distinct student population (i.e., only K-2). However, in both cases the core issue is missing assessment data because the statewide assessment is not administered prior to grade three. In other words, regardless of how we categorize the exceptionality, the factor driving the need for a unique solution is the same (missing data).

Because the accountability issues related to alternative schools are significant for many states, we devote the entirety of section two to this topic. In section three, we broadly discuss some common exceptional circumstances and how they might be addressed.



The Solution Space

Before we address the specific issues and options associated with alternative schools and other types of exceptionality, it is important to better understand the range of potential approaches in the solution space. To be clear, we do not assume that every exceptional school or circumstance requires a complete system redesign; nor do we assume the nature and scope of revisions in response to these exceptionalities should be similar across systems even when trying to address the same issue. Context and priorities matter.

As with any design problem, it is critical to start with an understanding of the specific problem(s) we want to solve and the conditions/constraints that inform alternatives. Only by understanding the context, priorities, and constraints can one identify potential solutions that are likely to be feasible and effective. Following, we list some guiding questions that can potentially help inform the solution space.

To what extent should systems support the same interpretations about school performance?

Whether the state is trying to design a distinct accountability system for alternative schools or deal with exceptional circumstances within a single system, it is important to understand whether the solution must produce outcomes at the indicator or overall level that are intended to be interpreted in a similar manner. Broadly, this may be interpreted as comparability, which is further addressed by subsequent questions about rigor and standardization.

When systems are comparable, one can meaningfully interpret performance for any school with respect to the ‘traditional’ or reference system. In a highly comparable system, the scores and performance ratings may use the same scale and language as the reference system and these outcomes are intended to be interpreted in the same manner. Systems not intended to be comparable may rely on distinct scores and performance descriptions with no explicit relationship to the reference system. There is a vast middle ground on the continuum of comparability that may include features such as:

- Some elements or indicators are intended to be comparable; others are distinct (e.g., academic growth is the same, but ‘readiness’ is different).
- Systems produce distinct performance categories, some or all of which are linked to the reference system (e.g., ‘adequate’ on one system corresponds to ‘meets expectations’ on another).

The desired or required degree of comparability to the reference system is based on a number of factors such as: federal regulations, state policy priorities, state context, and state capacity to provide supports and interventions. There are also different strengths and limitations of system design, depending upon the degree of comparability. For example, highly comparable systems may allow for similar scores among schools, but not meaningfully differentiate school quality between traditional schools and exceptional schools. On the other hand, weakly comparable systems may utilize different indicators of school quality and student success, which could meaningfully differentiate school quality, but appear publicly and politically as not holding all schools to the same accountability yardstick. This relates directly to the next guiding question that can potentially help narrow the solution space.



Should systems have similar levels of rigor?

We use rigor to refer to the relative challenge or demand reflected in performance expectations. As noted, rigor influences the manner in which outcomes can be interpreted similarly. However, it is possible to design an alternative accountability system intended to support different interpretations, each of which has a relatively consistent level of rigor. Conversely, it is possible to design an alternative accountability system intended to support the same interpretations, but at a different level of rigor (e.g., simply lowering or raising the bar within the same system). We believe the latter case is rarely a desired objective.

A claim that two distinct accountability systems demonstrate similar levels of rigor may be supported in different ways. The most obvious approach is to design for equivalence by using the same indicators and performance criteria within both systems. For example, both systems may reward schools for having 70% or more of the students achieve proficiency on state tests. Alternatively, claims of similar rigor may refer to the relative challenge in a normative sense. For example, 10% of schools may receive the highest classification in both the reference system and in the alternative accountability system. Still another approach is to establish different expectations for rigor based on expert judgment. Specifically, experts would be asked to identify the level of performance in the alternative system that represents a similar degree of rigor as the expectations defined for the reference system (Domaleski, D’Brot, & Keng, 2018). This may be particularly relevant when indicators or other system characteristics are intentionally different and the distribution of school performance is not assumed to be equivalent.

Is standardization important?

In the context of accountability systems, standardization refers to the consistent application of uniform rules to produce school accountability results. It may be helpful to think of standardization as supporting comparability within a system, which is necessary but not sufficient to support comparability between or among systems.

Like other dimensions, standardization exists on a continuum. When the same exact indicators and business rules are used consistently for all schools, a system is highly standardized. Another approach to standardization is to allow different indicators, but use the same general rules to evaluate performance. For example, many school accountability systems permit some “or” conditions, such as rewarding credit for Advanced Placement (AP) courses or dual enrollment courses. Strictly speaking, the indicators are different but may be judged to be sufficiently similar to support comparability. These claims are bolstered when the rules for awarding credit are based on standardized procedures and criteria (e.g., similar course content and similar qualifying assessment scores). A less standardized approach may involve greater selection from among a broad “menu” of more distinct indicators. For example, an approach in which credit is available for a wide-range of internship or work experiences may be less standardized. The guiding principle for such an approach may be that any qualifying experience should meet minimum criteria for acceptability, even if claims about strict standardization for all attributes and features are not feasible.



How are consequences addressed?

A key purpose of school accountability is to specify the consequences associated with system outcomes in order to incentivize schools to attain or maintain a level of performance that meets the state's expectations. Consequences may take the form of supports and interventions (e.g., monitoring, professional development, etc.) or rewards (e.g., public recognition, relaxing state requirements, etc.). At the early stages of identifying the solution space, we argue that it is critical to understand if/how consequences should differ for exceptional schools or circumstances in comparison to traditional schools or circumstances.

If the information provided is largely useful, but the consequences are not reasonable or effective, the solution may be to focus primarily on changing the system of supports and interventions. In other cases, the incentives, information, and consequences associated with the reference system may be poorly suited to help certain schools improve, in which case more comprehensive changes will be required.

Organization

In section two of this paper, we focus specifically on exceptionality as it relates to alternative schools. We devote considerable attention to this topic, given that it is a prominent issue for state accountability systems, and because the solutions often involve a more extensive process. We start section two by addressing the characteristics and features of alternative schools and then explain why these elements make the use of a state's standardized model so challenging. Subsequently, we describe a three-phase process for developing accountability solutions for alternative schools.

In section three, we address exceptional circumstances. We describe the types of circumstances that can impede the application of standardized accountability models. Then, we suggest a typology for potential solutions, providing examples within each category.

Section Two

Designing Alternative School Accountability Systems: What, Why, and How?

As noted in section one, exceptional schools are those with distinct student populations, missions, or characteristics. There are many types of exceptional schools—some publicly funded and some privately funded. We focus this section on public alternative schools as one common type from which key characteristics and considerations could be extrapolated to other exceptional schools, as applicable.

Characteristics of Alternative Schools: The What

Alternative schools represent a broad range of schools that primarily serve students who have not been successful in a traditional school environment (Carver & Lewis, 2010). The National Center for Education Statistics (NCES) defines an alternative education school as “A public elementary/secondary school that (1) addresses needs of students that typically cannot be met in



a regular school, (2) provides nontraditional education, (3) serves as an adjunct to a regular school, or (4) falls outside the categories of regular, special education, or vocational education” (Keaton, 2012, p. B-1). According to the common core of data, as of the 2010-11 school year, about 6% of public elementary and secondary schools in the United States were considered alternative (N = ~6,200), which represents about 560,000 students (Keaton, 2012). These schools and programs are designed to address the needs of students that typically cannot be met in traditional schools (Porowski, O’Conner, & Luo, 2014). As a result, alternative schools and programs that receive accountability ratings typically perform poorly in traditional school accountability systems that feature and heavily weight indicators of academic achievement, chronic absenteeism, and/or traditional measures of college and career readiness (e.g., graduation rate, Advanced Placement/International Baccalaureate participation, and/or meeting College and Career Ready benchmarks).

That said, alternative schools are not a homogenous group. Alternative schools differ from traditional schools as well as from each other. In fact, there is likely more variability among alternative schools due to the populations of students served than is typical among traditional schools. This is due in part to the fact that states have different policies, processes, and procedures for determining which schools are designated with alternative status (Porowski et al., 2014). Some states select/determine alternative status; in other states schools submit an application to the state for review and approval. Debate exists within and among states about which schools should ‘count’ as alternative schools. Our goal in this paper is not to dictate a set of rules for defining or designating alternative schools. Instead, we discuss some of the key features that differentiate alternative schools from traditional schools, motivating the design of alternative school accountability systems, while acknowledging the vast variability among alternative schools.

Alternative schools can be characterized in many ways. Porowski, O’Conner, and Luo (2014) argue from a scan of state definitions that there are four key aspects of alternative education: the target population, setting (e.g., within a school or in a standalone school), services offered, and structure (e.g., during or outside of school hours). Aron (2006) discusses many dimensions that could be used in the development of a typology for alternative schools, including: the target population; the program’s purpose or focus; the physical setting relative to regular schools or other institutions such as residential treatment or juvenile justice facilities; the educational focus or credential offered; the administrative home or sponsor; and how it is funded. Still others define alternative schools more broadly based upon the characteristics of students who generally attend, including those who are at risk of educational failure due to poor grades, truancy, disruptive behavior, pregnancy, or similar factors (Carver & Lewis, 2010).

One way to synthesize across these various definitions is to describe the characteristics of students served and the characteristics of the school that differ from traditional schools. Table 1 below contains one such typology. It summarizes the features that distinguish alternative schools from traditional schools and alternative schools from one another with respect to key student and school dimensions. The features in the right-hand column are intended to be neither exhaustive nor independent; however, different alternative schools may emphasize different features due to the diverse nature of the students served. This typology is summarized in the paragraphs that follow.



Table 1
One Possible Typology of Alternative Schools

Typology	Dimensions	Features
Characteristics of Students	Grade levels and ages served	<ul style="list-style-type: none"> ✓ High school students most common, followed by middle school and more rarely elementary school age students
	Target population	<ul style="list-style-type: none"> ✓ Students with behavioral problems (e.g., students who disrupt the classroom, students who commit severe disciplinary infractions) ✓ Students with academic problems (e.g., students with poor grades, low standardized test scores, students who are behind in school credits) ✓ Students at-risk of dropping out or not graduating on time (pregnancy, homelessness, drug or alcohol abuse, and physical or sexual abuse) ✓ Students in need of additional/supplemental services beyond those provided by traditional schools (e.g., students with disabilities) ✓ Students who have dropped out ✓ Students with truancy or attendance problems
Characteristics of Schools	Mission & Purpose	<ul style="list-style-type: none"> ✓ One type of alternative school focuses on students who need more individualized learning to succeed, or on dropouts wanting a diploma. Full instructional program offers students credits for graduation. Students may choose to attend. May be located within schools. ✓ One type of alternative school focuses on discipline and the purpose of reforming disruptive students or students with major behavioral issues. Students typically do not choose to attend, but are sent to the school for a specified time period or until behavioral requirements are met. Placement is typically short-term. ✓ One type of alternative school focuses on offering therapeutic settings for students with social and emotional problems. Students may or may not choose to attend, but placement may be short-term or long-term depending on the severity of student needs.
	Setting	<ul style="list-style-type: none"> ✓ Within regular schools ✓ Separate site or school facilities ✓ Online/distance learning



		<ul style="list-style-type: none"> ✓ Residential facilities ✓ Juvenile justice facilities
	Services Offered	<ul style="list-style-type: none"> ✓ Regular academic instruction ✓ Remedial education ✓ Tutoring ✓ Academic support ✓ Behavioral services/counseling ✓ Social skills and support services (e.g., social/life skills instruction) ✓ Career education (e.g., employability skills, job training, vocational education, career counseling) ✓ Transportation ✓ Daycare or childcare ✓ Health services ✓ Special education services ✓ Substance abuse interventions ✓ Links to community partners
	Structure	<ul style="list-style-type: none"> ✓ Operate during school hours ✓ Operate during non-school hours, such as evenings, weekends, and summers ✓ Small size and small classes ✓ Personalized, whole-student approach ✓ Student to staff ratio

Unique Characteristics of Students

Primary among the unique features of alternative schools is the student populations served. Previous surveys of alternative schools report alternative schools are most common for high school students, followed by middle school students, and least common is elementary school students (Carver & Lewis, 2010; Keaton, 2012). The target population for alternative schools is primarily students at risk of educational failure either due to poor academic performance, disruptive behavior, truancy, disability, and/or social or emotional needs.

Unique Characteristics of Schools

An alternative school’s specified mission/purpose, setting, services offered, and structure should flow directly from the student population(s) served. In other words, the way a school operates, how the environment and instruction are structured, and the types of services provided should follow from the characteristics of students served. Alternative schools should be doing something different for students and operating distinctly because of the unique challenges and opportunities faced by their students. For example, alternative schools that focus on serving students who are not succeeding in a traditional school because of disruptive behavior will make distinct choices about school setting (separate site or facility), services offered (behavioral services/counseling, remedial education, academic support, etc.), and structure (operate within



schools hours, but may also offer wraparound services over the summer and weekends) in order to best meet the needs of their target population. In comparison, alternative schools that primarily serve students who need more individualized instruction because of a disability, pregnancy, truancy issue, etc. will cater their school setting (within regular schools, distance/online learning), services offered (regular academic instruction, career education, etc.), and structure (operate within school hours, personalized instruction, small class sizes) to meet those students' needs. The alternative school environment and instruction should be tailored to student needs with respect to the degree and type of flexibility necessary to support student success.

Common Data Challenges for Alternative Schools

Some common data challenges of applying a one-size-fits-all school accountability model to alternative schools are the availability of data on these students and the small sizes of alternative schools. For example, alternative schools often serve a highly transient student population with students who may transfer in and out of schools within a school year. This poses limits and challenges to the availability of information that can be collected on each student, and presents difficulty identifying which school to attribute each student's data. Additionally, most alternative schools serve small numbers of students as a function of their unique mission, goals, and student population in comparison to traditional schools. Alternative schools are often small because they are designed to ensure individualized attention and intensive remediation and support to students who are far behind grade level and/or present with significant learning needs.

Rationale for Alternative School Accountability Systems: The Why

In most states, alternative schools perform very poorly on the indicators measured by the traditional accountability system in comparison to traditional schools. This is not unexpected given the unique characteristics of the students served and the fact that students typically attend an alternative school because they were not well-served in a traditional school setting.

Low alternative school performance on indicators measured by the traditional accountability system can be interpreted in one of two ways: (1) it is a true indicator of performance; or (2) the system is not well-designed for alternative schools. For example, if most of the alternative schools in a state receive an 'F' using a letter-grade system, is it because all of these schools are low performing or because the traditional school system is not well-designed for these distinctive schools? If many alternative high schools are designated as in need of comprehensive supports and interventions because they perform in the bottom 5% of schools in the state, is it because those schools are failing the students they serve or because those schools have a distinctive mission to help students who are well-below grade level, highly transient, and/or chronically absent?

Increasingly, many educators, school/district leaders, and policymakers have realized the latter interpretation (the system is not well-designed for alternative schools) is a more persuasive position, and that the purposes of school accountability systems are often thwarted (and even skewed) when applied to alternative schools. The accountability model applied should better reflect the educational programming provided by alternative schools because of the students they serve. Ultimately, an accountability system that doesn't fit fails to differentiate low and high



performing alternative schools, provide useful feedback that informs improvement, and incentivize the right actions.

The goal when designing an alternative school accountability system, therefore, is to better reflect the unique characteristics of this subset of exceptional schools and the students they serve. This alternative system then provides relevant and meaningful accountability aligned with the prioritized purpose(s) of the system (e.g., motivate system change, provide useful data for decision-making, align system components, and signal system accountability to stakeholders).

One argument that could be levied against the establishment of alternative school accountability systems is that the intent of such systems is to reduce rigor or lower the standards of performance for these schools. States or districts could ‘game the system’ and intentionally designate schools as alternative and funnel at-risk students to those schools in order to bolster the traditional state school accountability report card or the overall ratings of other schools in the district. This is an important consideration and a possible unintended consequence that must be considered throughout the system design.

This argument assumes, however, that there are no other safeguards in place to ensure that states and districts have a clear process for assigning alternative status, are transparent with which schools are designated as alternative, and provide reasons/justifications for such designations. Moreover, this assumes that the rigor of the alternative system will be lower than that of the reference system. We do not assume this is the case. In fact, in our experience, states that have pursued alternative systems have embraced rigorous standards of performance. The approach of many states can be described as: different system, but same high expectations.

Another caution in this vein is that just as we are arguing against a one size fits all model for school accountability, it is important not to assume a one size fits all model for highly diverse alternative schools. Design follows purpose, priorities, and goals. In the next section of this paper, we show how a state could use a theory of action approach to design a state alternative school accountability system that recognizes the diversity among alternative schools within a state and also recognizes the differences between alternative and traditional schools.

The Design of Alternative School Accountability Systems: The How

As suggested in section one of this document, there are several considerations that inform how one identifies an appropriate solution for addressing exceptionality in state accountability systems. When it comes to the development of an alternative school accountability system, a state’s approach to these factors determines how and to what degree the new system may differ from the reference model. To illustrate, Table 2 highlights some of the dimensions on which a state’s alternative school accountability system might be similar to, or different from, the traditional school accountability model.



Table 2

Dimensions of Difference for Traditional and Alternative School Accountability Models

Dimension	Similar	Different
Selection of Indicators	Both systems include the same set of performance indicators	Alternative school system identifies additional or different indicators
Performance Standards	Alternative schools are held to the same expectations for performance on common indicators	Different expectations for performance are defined for alternative schools on one or more common indicators (e.g., achievement)
Flexibility & Choice	Overall school ratings or performance determinations are based on a common set of state-defined indicators that apply to all schools.	Alternative school leaders can select from a menu of state-specified indicators those which would best support interpretations about their school for use in establishing an overall rating.
Weighting and Aggregation	Common indicators receive the same weight across both systems. The methods used to aggregate results across indicators are the same (e.g., index, decision matrix)	Procedures used for aggregation and/or the relative emphasis afforded to common indicators differ from that represented in the traditional school model.
Reporting of Results	Results are reported using the same labels and performance categories represented in the traditional system.	Alternative school results are reported on a different metric and/or using a unique set of performance categories.

How these dimensions are represented in the design of a state’s alternative school accountability system depends on a variety of factors including the state’s goals and priorities, and contextual elements that serve to constrain or dictate the system design (e.g., stakeholder expectations, legislative requirements, and resources to support implementation).

In the subsections that follow, we outline a three-phase process that state education agencies (SEA) can use to articulate a vision and design for an alternative school accountability system. The process uses a theory of action framework to establish a rationale for the system design, facilitate collaboration with stakeholders, and provide a roadmap for system evaluation.

Phases of the Design Process

In its simplest form, a theory of action is an argument that explains how the design of a system will meet its specified goals. Inherent in this argument is the designers’ hypothesis for how the goals will be achieved, the outcomes that signal progress toward or attainment of the goals, and the assumptions that must hold in order for the system to function as intended. Like other complex systems, school accountability models are made up of multiple parts, each of which is intentionally identified and designed to play a unique role. Unfortunately, the more complex the system the more difficult it is to control and anticipate the factors that can prevent it from

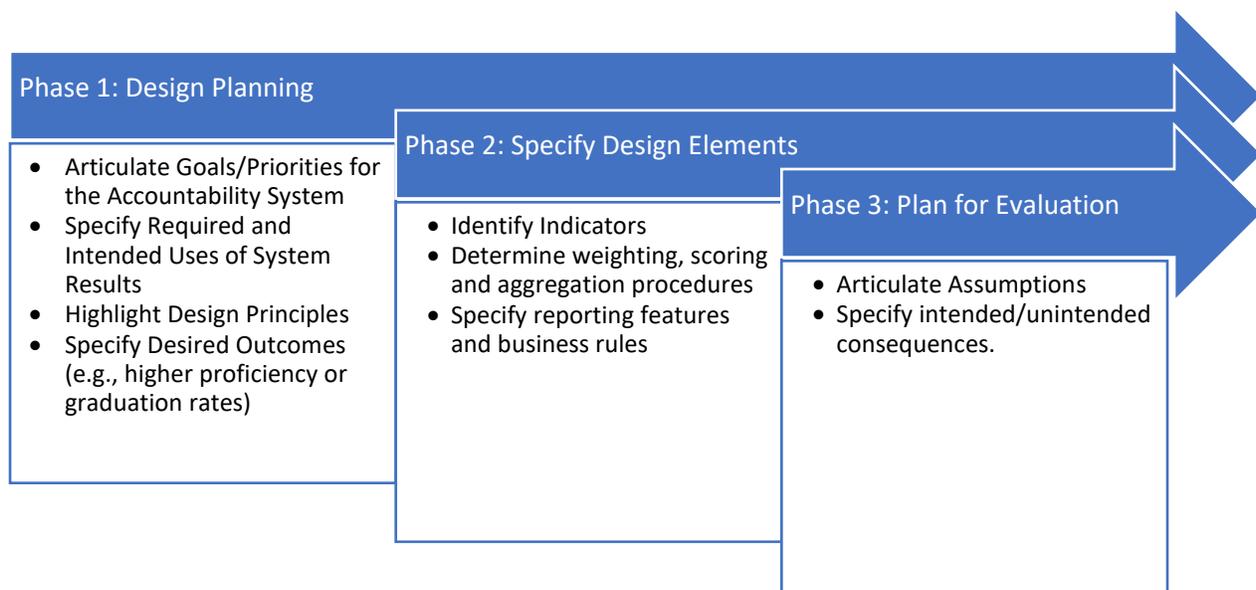




working as intended. This is especially the case with state-designed school accountability systems which are often developed by multiple teams or divisions working independently. A well-articulated theory of action requires the rationale, requirements and assumptions underlying each design decision to be explicitly stated, so that potential issues can be identified and addressed during the system design.

To structure our discussion around the design of alternative school accountability systems we refer to Figure 2 which outlines nine design tasks organized in three levels. Each level represents a different phase of the design process and in combination supports the generation of a comprehensive theory of action.

Figure 2
Phases of the Assessment System Design Process



While the design process is hierarchical and linear (e.g., design elements cannot be articulated without first engaging in design planning), it is also iterative in that early decisions may be revisited and revised at any point in the process.

In Table 3 below we highlight some of the key questions to be addressed in each phase. We discuss each phase in turn, focusing on how and why states' responses to these questions might vary. The intent of this discussion is *not* to lay out a set of design solutions that states can select from in a check-box fashion. Rather, we attempt to highlight the process and thinking necessary to design a coherent accountability system that aligns with a state's unique priorities and goals. While the questions in Table 3 refer to alternative schools, this process and set of questions can be generalized to support the development, evaluation or modification of any accountability system, including those discussed in Part 3 of this paper.

While the alternative schools within a state may differ greatly, as mentioned in the previous section, *we are not suggesting a separate accountability system be developed for each*



alternative school. We do recommend, however, that appropriate stakeholders be engaged in all phases of the design process so that any concerns about the appropriateness of the system for all alternative schools can be identified and addressed. At that time, decisions can be made about how/if the model needs to be revised to support one or more alternative schools.

Since the design elements and plan for evaluation flow from Phase 1, the bulk of our discussion is focused on this level.

Table 3
Key Questions by Design Phase

Phase 1: Design Planning	
Key Questions	Key Stakeholders
<p>Vision:</p> <ul style="list-style-type: none"> ● What are the goals of the accountability system? Which goals, if any, are unique to the alternative school accountability system? ● What conditions, information, and supports are needed to achieve the state’s goals? <p>Intended Uses:</p> <ul style="list-style-type: none"> ● How are the results intended to be used and interpreted? ● What claims must be supported to use the results as intended? <p>Desired Outcomes:</p> <ul style="list-style-type: none"> ● How would you know if the alternative school accountability system is working as intended? <ul style="list-style-type: none"> ○ What student outcomes would you expect to see? ○ In what ways might adult practice or policies change? <p>Design Principles:</p> <ul style="list-style-type: none"> ● What values, principles or rules should be represented in the system design? ● What design features must hold in order for the alternative accountability system to produce results that can be used and interpreted as intended. 	<p>State Education Agency</p> <p>District and Alternative School Leaders</p>



Phase 2: Design Elements	
Key Questions	Key Stakeholders
<p>Indicators:</p> <ul style="list-style-type: none"> • What indicators (e.g., scores, measures, ratings) should be included in the alternative accountability system and why? • What, if any, indicators should be common to those defined in the traditional accountability system and why? <p>Weighting and Aggregation:</p> <ul style="list-style-type: none"> • What weights will be assigned to each indicator and why? • How will the indicators be combined or aggregated (e.g., decision matrix or index score)? <p>Reporting and Use of System Results:</p> <ul style="list-style-type: none"> • What results will be reported to schools (e.g., indicator scores, overall school rating) and the public? • What business rules must be in place to ensure reported results are accurate, reliable and can be used as intended (e.g., definition of full academic year)? • What supports/tools/guidance needs to be developed or provided to ensure the system is implemented with fidelity and results are used as intended? 	<p>District and Alternative School Leaders</p> <p>Alternative School Teachers;</p> <p>Parents and caregivers of alternative school students</p>
Phase 3: Planning for Evaluation	
Key Questions	Key Stakeholders
<p>Assumptions</p> <ul style="list-style-type: none"> • What conditions must hold in order for the alternative accountability system to have the desired impact? <p>Consequences</p> <ul style="list-style-type: none"> • What are the potential intended and unintended consequences of implementing the proposed alternative accountability system? How might these consequences be supported, monitored or mitigated, as appropriate? 	<p>District and Alternative School Leaders;</p> <p>Alternative School Teachers</p>



Phase 1: Design Planning

The purpose of the design planning phase is to ensure the vision for the alternative school accountability system and all factors that influence, constrain or inform that vision are clear. For this reason, alternative school leaders must be included in this discussion. Not only do they understand the characteristics of these schools and the students they serve, they are in the best position to determine the conditions necessary to meet the state's goals and how they can be incentivized in the system design.

Goals & Vision

The first step in the design planning phase is to clarify the goals of the accountability system and sketch out how they are likely to be attained. Goals reflect a state's overarching beliefs regarding the purpose of an accountability system and the role it should play to support student learning. Since these beliefs are not likely to differ for alternative and traditional schools, several (or all) goals may be the same for both systems. Some examples of accountability system goals include:

- Provide appropriate and targeted support to the lowest performing schools.
- Incentivize and support the attainment of skills and credentials that promote post-secondary success.
- Provide a valid indicator of overall school quality and performance within and across years.
- Provide schools and stakeholders (community, state leaders, policy makers, teachers and schools/district leaders) with accurate information that informs local improvement efforts.
- Improve collaboration between alternative schools, post-secondary institutions, and local employers.
- Increase local support for alternative schools.

While the last two goals are unique to alternative schools, in theory a state could identify a common set of state goals and still develop two completely unique school accountability systems. Ultimately, it is the state's goals in combination with its beliefs as to how those goals will be met that influences the system design. For example, if an SEA believes small teacher-to-student ratios and personalized learning are the key to supporting student success in alternative schools, it may design the alternative school accountability system to incentivize small class sizes and opportunities for individualized support. Similarly, a state that believes flexibility of indicators is necessary to obtain valid measures of alternative school performance might design a system that lets schools identify or choose among multiple measures of school quality. These practices are not typically represented in traditional school accountability models even if the goals (e.g., support student success, establish a valid measure of school quality) are the same.

Outcomes

Outcomes are the observable, measurable changes you would expect to see if the accountability system is working as intended. Outcomes operationalize the state's goals and clarify the state's priorities in a manner that stakeholders understand and schools can monitor. Many of the desired outcomes for alternative and traditional school accountability models will be the same (e.g., increased rates of academic achievement, growth, credential attainment, and other measures of college and career readiness); however, outcomes that better align to the mission and goals of alternative schools may also be defined. These include things such as: more students graduating with internship experience and professional certifications; decreasing dropout rates; higher rates



of teacher retention; and positive feedback from alternative school leaders regarding the quality and utility of accountability system results.

Uses

Another aspect of the design planning phase is clarifying how the results of the alternative school accountability system are intended to be used. This impacts both the format the results will take and the interpretations they must support. For example, if a state’s accountability system results will be used to identify the lowest performing 5% of alternative schools the system must produce an overall index score or rating that supports this activity. Similarly, if the results will be used to inform decisions about the quality of a new program or initiative the system must include measures that will be positively impacted if the program/initiative worked as intended.

Designers must also clearly define how information at different levels of the accountability system will be used (e.g., overall scores or rating, indicator scores or ratings, individual measures). For example, if a state intends to aggregate and report a growth rating for all schools in the state (e.g., low, medium, or high) the labels used to report growth under the alternative accountability system should mirror those defined within the traditional school model. If the full array of intended uses is not identified early on, indicators may not be designed and reported in a way that supports the state’s goals.

Design Principles

If the goals represent the intended destination on a roadmap, the design principles guide the nature and manner of the route. They are the values, priorities and constraints that need to be considered when identifying and defining the elements of the accountability system. Examples of common design principles for school accountability systems include: *equity*—ensuring equitable outcomes for all schools; *rigor*—holding all schools to rigorous expectations for performance; and *comparability*—providing results that allow for the performance of schools to be compared.

An alternative school accountability system may wish to reflect these principles in addition to those believed necessary to support the success of alternative schools (e.g., flexibility, choice, and multiple pathways to success) or garner support for the alternative school accountability system (e.g., transparency). Since some principles can work against each other (e.g., flexibility and comparability) stakeholders must identify which principles should receive the highest priority in the system design.

Phase 2: Specify Design Elements

The purpose of Phase 2 is to determine the indicators, procedures and outputs that need to be in place in order to meet the priorities and expectations defined for the system design. Each activity is briefly discussed below and examples are provided to demonstrate how they might be influenced by Phase 1 decisions.

Select Indicators

Indicators are measures or ratings that:

- Provide information about progress toward or the attainment of desired outcomes;
- Incentivize conditions or interactions believed to support the attainment of desired outcomes; and
- Communicate the characteristics of a quality school.



A state should select and define indicators for its alternative school accountability system that meet one or more of these conditions and directly reflect the priorities, expectations and any constraints identified in the Design Planning phase. For example, if a state believes that a positive and nurturing school climate is a necessary condition to support student success, a measure reflecting student perceptions of school climate may be included in the system. Similarly, a state may include an indicator of career and technical education (CTE) course representation (i.e., the number of CTE programs of study offered within a school) as a way to incentivize alternative schools to provide students with more opportunities to earn a career/industry certification or credential.

Due to the unique mission of alternative schools, less common indicators (e.g., school climate, student engagement, participation in community service, establishing citizenship) are often considered for inclusion in alternative school accountability models. While this may be reasonable, a state must carefully consider how the results will be used as these indicators are often less reliable and susceptible to gaming. Other factors to consider include the availability of data and the time and effort necessary to collect the data of interest. For example, a career-readiness indicator that requires gathering and scoring portfolios of student work could put an undue burden on educators and take away from instructional time.

Indicators are often proposed during Phase 1 discussions. In Phase 2, decisions are made about how indicators should be scored, aggregated and reported to meet the goals of the system.

Weighting, Scoring & Aggregation

Weighting is the process of distributing emphasis to each indicator in the accountability system. A variety of factors can influence decisions about how indicators should be weighted, including: perceived significance, reliability, technical quality, and desired impact on the overall rating and/or stakeholders' perceptions. Unfortunately, in many cases the indicator believed to be the most relevant may also be hard to collect, unreliable, or susceptible to corruption. This can be especially problematic in alternative school accountability systems where there may be a desire to use indicators based on surveys, self-report measures, observations and locally defined artifacts of student performance. States must carefully consider the role and purpose of each indicator in the system and balance these against any technical concerns. If the primary purpose of an indicator is to incentivize action (e.g., quarterly student/teacher conferences) or signal a key priority (e.g., safety) including the indicator in the system may be enough to have the intended impact regardless of assigned weight.

In many cases, raw indicator measures (e.g., average school climate scores; drop-out rates) may need to be scored to support aggregation and/or the intended interpretation and use of results. Scoring involves transforming the raw indicator measure to a metric or scale having desired properties. In some cases, indicators are scored in multiple ways to support different purposes. For example, a state may convert each measure in its system to a 0-100 scale in order to establish an overall weighted index score and also establish standards that allow for performance on each indicator to be classified as meeting/not meeting expectations.

Reliability may also influence how results are scored. For example, less reliable measures can be transformed to performance categories to increase confidence in the results and mitigate misinterpretations.



There are a variety of procedures that can be used to aggregate data across a series of indicators (e.g., data matrix, index, and dashboard). Each procedure reflects a different priority and has an array of associated pros and cons (Reyna, 2017). While a discussion of these procedures is outside the scope of this paper, it should be apparent from previous sections that several factors (e.g., policy, practical, technical) can influence the procedures selected to combine and report results.

Reporting & Business Rules

The final component of Phase 2 involves clarifying how and what will be reported and the business rules necessary to ensure the system provides for accurate results. Alternative school leaders should play a large role in this discussion by clarifying the type of information needed and any data issues/concerns that should be addressed in the business rules. If desired reporting features are not collected as part of the design process, the results of the system may not be used by stakeholders as intended. Feedback related to reporting should include such things as the level of granularity at which information should be reported, the types of interpretations the results should support, and the data displays that will be most useful (e.g., trend data) to support school improvement efforts.

Business rules refer to a comprehensive set of requirements, definitions and criteria that clarify how each element of the accountability system is calculated and reported (e.g., eligibility requirements, N-counts, definition of a full academic year, data calculations and rounding rules). Because they dictate how and what information is presented to stakeholders, business rules can have a significant impact on the utility of the accountability system and the validity of the inferences it is intended to support. While all accountability systems have business rules, due to the data challenges discussed earlier (e.g., small N-sizes, student mobility), business rules for an alternative school accountability system must be carefully considered throughout the design process.

Phase 3: Preparing for Evaluation

One of the primary benefits of defining a theory of action is that it informs the development of a plan for system evaluation. Phase 3 of the design process requires identification of the assumptions that must hold in order for the system to work as intended and any potential unintended consequences. Common assumptions include expectations related to:

- The availability of required resources and data;
- The technical characteristics of indicators and their relationship to each other as well as school-related factors;
- Appropriate data use and interpretation by stakeholders; and
- The frequency, utility and fidelity of intended interactions between stakeholders (e.g., students, educators, principals, parents).

For example, an indicator of school climate may be seen as a way to improve school quality by highlighting key areas of concern for school leaders, but this assumes there will be adequate participation by students and that school leaders know how to use the results to improve school climate. Similarly, an indicator of AP course-taking will not have a positive impact on participation or serve to differentiate school performance if alternative schools do not have the resources to offer these classes. To the extent possible, it is important to identify problematic



assumptions up front so decisions can be made about how/if the design should be modified prior to implementation.

It follows from these examples that technical advisors and stakeholders play a large role in both identifying assumptions and providing feedback on how/if they are likely to hold. An insider's perspective can also highlight potential unintended consequences and strategies for mitigation. For example, the development of student work portfolios, previously discussed, may frustrate educators and lead to increased rates of educator attrition in alternative schools. This may be mitigated, in part, by limiting the number of artifacts required and/or providing a user-friendly system to submit and score student work. Key stakeholders are critical partners in helping define what is reasonable and feasible.

Once articulated, a formal plan for system evaluation which highlights the evidence necessary to evaluate assumptions and monitor potential negative consequences can be established. A framework that can be used to support the evaluation process is discussed in a previous resource distributed by CCSSO (Hall, Domaleski, Russell, & Pinsonneault, 2016).

While this section addressed alternative school accountability system design, the process can be applied to a broader array of exceptional schools. In the next section, we discuss procedures for dealing with exceptional circumstances in accountability. If these circumstances preclude the traditional model from being implemented with fidelity a process similar to that defined in this section can be used to determine how/if the model should be modified to provide stakeholders with accurate/useful information about the performance of impacted schools.

Section Three

Exceptional Circumstances in Accountability

As noted in the introduction, exceptional circumstances refer to any condition that may impede or complicate the application of a standardized school accountability model. We distinguish exceptionality activated by circumstances, from approaches for exceptional schools as addressed in the previous section and depicted in Figure 1. We make this distinction because very often the goal of addressing exceptional circumstances is not to initiate a distinct policy solution, guided by different principles or values. Rather, the primary issue to be addressed is how to provide credible scores for use within an existing system. We caution that changes, even relatively minor variations, may inhibit the ability to support the same meaning, interpretation, and uses of accountability results within year or across years.

Implications of Exceptional Circumstances

Exceptional circumstances can arise from school specific factors or from events, such as testing irregularities or emergency conditions. School-specific factors include schools with atypical grade configurations or other characteristics that may impede application of standardized accountability models. A school serving grades K-2 is a common example, since most states administer statewide assessments for accountability beginning in grade three. However, atypical grade configurations can also lead to different indicators or dissimilar sources of influence among the same indicators. Consider a school covering grades K-12. That school will likely have



different indicators than a school covering grade spans of K-5, 6-8, or 9-12 alone. The combination of indicators and weights can create an aggregate outcome that is dissimilar from the results that would have been produced if any or all of these grade spans were considered individually.

There are certainly other school factors or characteristics apart from atypical grade configurations that can lead to exceptional circumstances in accountability. For example, schools with distinct student populations (e.g., incarcerated students) or missions (e.g., some charter schools) may fall into this category. In general, we expect these conditions will either inhibit a standardized implementation of the general model or call for a broader scrutiny of accountability design, following a process addressed in section two of this brief.

Another source of exceptional circumstances may be a disruptive event. For example, the widespread school closures and suspension of state testing due to the COVID-19 pandemic introduced a disruption to accountability systems of unprecedented scale. In fact, the disruption is likely to have a broad and multi-year impact (Domaleski, 2020). Other disruptive events can occur due to factors such as natural disasters or some administrative disturbance, such as irregular testing. Whatever the source, these events can either changes the nature or interpretation of data collected (e.g. are the test scores trustworthy?) or may lead to missing information altogether.

Whether due to school specific factors or disruptive events, exceptional circumstances often lead to missing or dissimilar indicators in an otherwise standardized model.

Missing Indicators

Exceptional circumstances may prohibit the use of one or more indicators in the model. Obviously, insufficient N-size is at the heart of any missing indicator issue, but this can occur for different reasons. For example, consider the following list of potential conditions leading to one or more missing indicators:

- A school serving a generally homogeneous population has sufficient N-size to report all indicators at the school level, but not for several student groups such as certain racial/ethnic groups or students who are economically disadvantaged;
- A school has several missing indicators at the school and group level because it is a very small school;
- A school has an unusual grade configuration such that it includes no grades that administer a statewide accountability test;
- A school is missing assessment data due to sustained closure or refusals to test in one or more grades and/or content areas; and/or
- A school is missing assessment data due to testing irregularities (e.g., a failure with the computer-based testing platform) or testing improprieties (e.g., scores invalidated due to cheating) that led to test invalidations.

These examples illustrate that various conditions can impact what data are missing and why they are missing. Naturally, these conditions influence the solutions thought to be more or less promising.



Dissimilar Indicators

Another result of exceptional circumstances occurs when a school has a dissimilar or different indicator in comparison with the ‘standard’ accountability system. As with missing indicators, there may be multiple reasons why this condition occurs. Two relatively common examples are presented below.

First, schools or districts may have received permission to use one or more alternative measures such as through a waiver, pilot program, or other source of flexibility. A prominent example is the flexibility provided in ESSA for states to permit Local Education Agencies (LEA) to administer a locally-selected national high school assessment in lieu of the state test in high school. This is widely interpreted to provide a pathway for LEAs to give the ACT or SAT, but could include other national tests used by institutions of higher education for entrance or placement (Domaleski & Gong, 2017). In fact, this flexibility may lead to multiple schools using multiple dissimilar assessments in the state accountability model. Regardless, any time flexibility is granted to use a dissimilar measure in an accountability system, it impacts all indicators that are influenced by that measure (e.g., achievement, academic growth, achievement gaps, etc.).

Dissimilar indicators may also be introduced when there is uneven availability of an indicator for schools statewide. For example, some accountability systems offer a choice or ‘menu’ of options to satisfy a criterion. This is not uncommon for indicators in categories such as ‘college/career readiness,’ which may list multiple outcomes deemed suitable signals of readiness (e.g., AP credit, IB credit, dual/joint enrollment credit, etc.). For schools that do not have access to some or all options on the list, this uneven availability may raise questions about dissimilar indicators.

Strategies for Addressing Exceptionality

To determine an appropriate approach for addressing exceptional circumstances in accountability, it is critical to understand the nature of the issue and to identify the criteria or principles for evaluating an effective solution. The considerations presented in section one of this brief may be useful to help identify the solution space for promising options. Further we suggest addressing the following questions to help investigate alternatives:

- Does the alternative promote practices that are consistent with the state’s policy priorities?
- Does the alternative support the state’s theory of action for promoting improved outcomes?
- Is it likely that the alternative will provoke unintended negative consequences?
- Does the alternative approach systematically advantage or disadvantage schools based on factors that should not be related to accountability outcomes (e.g., large or small schools do not attain favorable scores)?
- Is the alternative practicable? Can staff at the state, district, and/or school level implement the alternative as intended based on available resources and capacity?
- Have appropriate leaders, stakeholders, and technical advisors reviewed the alternative?
- Does the alternative yield comparable results?



Finally, as with any proposed design decision related to state accountability, it is important to consider how evidence can be collected and evaluated to determine if the alternative is working as intended and inform refinements as appropriate.

In this context, we suggest three general approaches for addressing exceptionality:

- *Technical Alternatives*: This category is intended to cover a wide range of alternatives related to calculating and aggregating indicators.
- *Design Alternatives*: This category addresses changes to the model such as altering indicators, expectations or business rules for assigning ratings.
- *Qualitative Alternatives*: This describes approaches that allow for decision making outside a standardized model for how schools are rated and/or held accountable.

The following table provides examples of alternatives in these categories and some prominent considerations associated with each one.

Table 4
Three General Approaches for Addressing Exceptionality

Alternative	Primary Purpose	Considerations
Technical Alternatives		
Adjust N-size requirement	Lowering N-size reduces the number of missing groups/indicators for smaller schools	Reducing N-size can inflate unreliability (results are less stable). Good solutions balance inclusion and reliability.
Multi-year averaging	Reduces the number of missing groups/indicators for smaller schools	Improves inclusion and stability but can create ‘lag’ between performance and outcomes. Can be combined with disjunctive rules such as “use average score or most recent score, whichever is best”.
Redistribute weights	Produce summative score or rating with missing indicator(s)	Redistribution can be designed to honor nominal influence of remaining indicators. However, if some indicators are more/less rigorous, redistribution can lead to uneven expectations.
Design Alternatives		
Assign results to another school (i.e., ‘routing’ or ‘pairing’)	Assigns a score or rating to a school that has substantial missing data to another school that is plausibly linked to that school (e.g., K-2 school gets the rating of 3-5 school most students attend)	Determining the routing rules can be complex when there are multiple ‘receiving’ schools. Raises questions about the source of influence for outcomes.
Adjust expectations for indicator(s) or overall rating	Address exceptional circumstance by raising or lowering rigor (e.g., different threshold ratings for K-12 school compared to 9-12 school)	Requires strong rationale and careful process to preserve comparability and interpretation of results for schools.



Add or change indicators	May address missing or uneven access to indicators (e.g., add additional career certifications)	Could inhibit comparability and may raise questions about different levels of rigor.
Qualitative Alternatives		
Implement appeals process	Addresses exceptional circumstances by giving school a process to appeal rating to a decision-making body.	Requires well-explicated process and criteria for hearing and adjudicating appeals. Can be very resource intensive.
Implement school review process	Addresses exceptional circumstances by <u>replacing</u> the standard accountability process with a decision-making body.	Requires well-explicated process and criteria for adjudicating appeals. Can be very resource intensive.
Policy adjustment to rating or consequences	Policy decision to deal with exception thought to influence a rating or consequence (e.g., legacy ratings issued as a ‘hold harmless’ due to testing system failure)	Requires strong rationale, criteria, and process. Impacts comparability.

Final Thoughts

We hope the ideas in this document help illuminate what types of exceptional schools or circumstances may require attention with respect to the traditional school accountability system. Our intent was to provide a rationale for why the “one size fits all” accountability model may not best serve the purposes of school accountability for all public schools within a state. Because there is no single correct approach for addressing exceptional schools or circumstances in accountability systems, we focused on guidance to inform the development and implementation of a process that attends to context, priorities, and constraints. We further argue that both general and alternative accountability systems are meant to be dynamic. That is, implementation should be accompanied by ongoing evaluation and analyses to assess the assumptions of the guiding theory of action. By so doing, the systems can be refined and improved over time to help realize their potential to improve student outcomes.



References

- Aron, L. Y. (2006). *An overview of alternative education*. Washington, DC: The Urban Institute.
- Carver, P. R., & Lewis, L. (2010). *Alternative schools and programs for public school students at risk of educational failure: 2007-08 (NCES 2010-026)*. Washington, DC: U.S. Department of Education, National Center for Education Statistics. Retrieved from <http://nces.ed.gov/pubsearch>
- Domaleski, C. (2020). *The outlook for school accountability after COVID-19*. Dover, NH: National Center for the Improvement of Educational Assessment. Retrieved from: <https://www.nciea.org/blog/essa/outlook-essa-school-accountability-after-covid-19>.
- Domaleski, C., D’Brot, J., & Keng, L. (2018). *Establishing performance standards for school accountability systems*. Washington, DC: Council of Chief State School Officers. Retrieved from https://www.nciea.org/sites/default/files/publications/Establishing Performance Standards for School Accountability Systems_0.pdf
- Domaleski, C. & Gong, B. (2017). *Implementing the locally selected high school assessment provision of the Every Student Succeeds Act*. Washington, DC: Council of Chief State School Officers. Retrieved from: https://www.nciea.org/sites/default/files/inline-files/Implementing-Locally-Selected-Assessment_4-17-17_0.pdf
- Hall, E., Domaleski, C., Russell, M., & Pinsonneault, L. (2016). *A framework to support accountability evaluation*. Retrieved from: http://www.nciea.org/publication_PDFs/A%20Framework%20to%20Support%20Accountability%20Evaluation.pdf
- Keaton, P. (2012). *Numbers and types of public elementary and secondary schools from the common core of data: School year 2010-11 (NCES 2012-325rev)*. Washington, DC: U.S. Department of Education, National Center for Education Statistics. Retrieved from <http://nces.ed.gov/pubsearch>
- Porowski, A., O’Conner, R., & Luo, J. L. (2014). *How do states define alternative education? (REL 2014-038)*. Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Mid-Atlantic. Retrieved from <http://ies.ed.gov/ncee/edlabs>
- Reyna, R. (2016). *Key issues in aggregating indicators for accountability determinations under ESSA*. Washington, DC: Council of Chief State School Officers. Retrieved from: <https://www.ccsso.org/sites/default/files/2017-10/KeyIssuesinAggregatingIndicators.pdf>
- Supovitz, J. (2009). Can high stakes testing leverage educational improvement? Prospects from the last decade of testing and accountability reform. *Journal of Educational Change*, 10(2–3), 211–227. <https://doi.org/10.1007/s10833-009-9105-2>