The Future of Assessment

White Paper 2
Introduction

Two years ago, High Resolves initiated the Future of Assessment project to explore innovative approaches for measuring the development of citizenship competencies in more robust ways than traditional assessment approaches. The project resulted in important breakthroughs in our thinking and a disruptive reframing that challenged the very notion of separating assessment from curriculum. The notion of embedding micro-assessments into learning experiences paved the way for delivering a new system for measuring impact that meets the needs of both learners and educators. During the same time, Got History has been working on an assessment project with a similar goal in mind: to increase demand for high-quality civic education in schools by making it possible to measure what matters.

The COVID-19 pandemic has intensified the conversation about the future of education, and on how to ensure that we are preparing our young people to thrive in and contribute to the world of today. With standardized tests on hold, there is the opportunity to reimagine how we assess learning so that our schools can help young people achieve the goals that truly matter for their success as human beings. Together, High Resolves and Got History have conducted extensive research over the past few months in the U.S. and Australia. We have formulated a pragmatic plan to deliver a proof of concept for this revolutionary new approach for assessment of citizenship competencies and other soft skills.

This White Paper, written in partnership with Fernande Raine of Got History, combines our evolving vision from the Future of Assessment project with a deepened sense of our current context. We hope it contributes to advancing the discourse on the future role of assessment in the transformation of education.
The Problem around Assessment

Turning Point for Reimagining Assessment

Our democracy is going through a time of significant challenges. The painful legacy of systemic racism and the disastrous effects of climate change have been exacerbated by the pandemic of COVID-19. Never has it been more urgent for us to focus our education system on preparing young people for a volatile, uncertain, complex, ambiguous and hyper-connected (VUCAH) world. We need to undertake urgent and collaborative efforts to orient our schools and universities towards achieving the goal of preparing competent global citizens who can ensure the wellbeing of themselves, their communities, the world and the planet as a whole.

For decades, education leaders across the globe have been engaged in deep and active debates about the purpose and key desirable outcomes of education as they tried to get a grasp on the challenges that lie ahead for humanity and did their best to imagine how best to prepare young people for success. They have produced powerful recommendations and policy documents outlining key priorities and learning frameworks for schools of today. Underlying all of these documents is the conclusion that the current system as it stands is not delivering the kind of preparation that our young people will need to thrive. Heightened levels of youth stress and anxiety, persistent economic inequality, disengagement of young people from democracy and a lack of hope are global indicators that we, the adults creating this education system, have not yet gotten it right.

If we want to ensure that our education systems meet this historic challenge, we must reimagine the field of assessment. As of today, there is no way of knowing whether schools are delivering what Danielle Allen has termed “participatory readiness” (Allen, 2016) and whether individual students are not only receiving the learning they need towards this goal, but are supported as individuals on their journey to becoming their full selves. None of the assessment tools in use to evaluate individual young people or schools speak to wellbeing and readiness to engage in a community. In the last few years, traditional knowledge-focused assessment models have been complimented with supplementary tools to capture social and emotional growth, but generally, teachers will agree that this is not enough. We need assessment models that will help orient our system towards preparing citizens for the world of today and tomorrow, and that can track whether individual students are growing along the dimensions that matter. These assessment tools will also play a critical role in helping schools find the interventions and curricular materials that can deliver the greatest impact at the lowest price. With a systematic approach to assessment, we will be able to answer some very important questions: what is citizenship education, who gets it, and how do we measure it today?
Citizenship Education: Definition and Adoption

Definition: Updated for how the world works today

Civics, civic learning and citizenship education are all terms used almost interchangeably with little attention to the nuance between them, and definitions vary. For the sake of this paper, we delineate them as follows:

- Civics is the study of the systems in which you live so that you can take informed action, including government structures, political parties, the press and civic action.
- Civic learning is the process of acquiring skills and knowledge relevant to being a member of a democracy.
- Citizenship education is the comprehensive package of learning that equips the next generation with the knowledge, tools, and dispositions needed to act for the wellbeing of themselves, their communities, and the planet.

Since citizenship education is the most systemic and broad term, it is the one we will use as we make a case for a new system of assessment so that democracy and our global community can thrive. With it we describe the full range of knowledge, skills and mindsets associated with global citizenship as they have recently evolved in the field.

The past 10 years have seen an explosion in the field of citizenship education, as educational entrepreneurs have responded to the shortcomings of our traditional schooling models with programs and interventions to supplement and complement the existing curricular offerings. Drawing on broad humanistic, democratic educational principles, these innovators—including High Resolves—have provided teaching tools that were usually geared towards humanities and social studies/history teachers and offered a new lens through which to approach educating young people. In 2018, a coalition of leading providers in the field agreed on a common framework to capture the broad set of mindsets, concepts and skills that are necessary for a young person to realize their own rights, enable those of others, and work in a community towards a collective vision of the future, described in Figure 1.

![Figure 1. Citizenship education as defined by Composer coalition](image-url)

This comprehensive framing of citizenship education goes far beyond what is commonly included in a civics curriculum. Citizenship education with this wide lens can “foster greater and more informed civic and political engagement among all youth, and in particular among youth from low-income families, youth of color, and immigrant youth. Indeed, high-quality approaches to civic learning also appear likely to promote the skills needed for the 21st century workplace, such as critical thinking, collaboration, creativity, initiative, and innovation. In addition, it can improve school climate and safety, lower schools’ dropout rates, and promote academic achievement aligned with the Common Core State Standards” (UCLA Project X, Gould et al, 2011). As an example, several studies have shown that an increased sense of belonging significantly enhances the ability of young people to absorb new learning. (Osterman, 2000).

Citizenship education not only activates the civic energy of individuals and enhances their ability to get along with others, it also provides the capacity of individuals to act with the collective orientation that has been proven necessary to build societies that can solve problems, survive epidemics and withstand sudden crises (Christakis, 2019; Harari, 2014). In a world of uncertainty, problems are interconnected and cannot be dealt with in siloed approaches, demanding citizens who are able to collaborate with a sense of shared values and a common vision for the future: this, too, is at the core of good citizenship education.
Adoption Challenges: Sidelined and Siloed

Despite the positive effects of citizenship education, school leaders broadly acknowledge that they are not set up to deliver its value to their students. According to a July 2018 survey of U.S. school leaders by Gallup, “three-quarters of district superintendents say preparing students to be engaged citizens is a challenge for schools,” representing a 24-point increase from the previous year (Gallup, 2018). In Australia, citizenship education faces similar challenges, perceived as less rigorous when compared to other academic disciplines (Henderson, 2016). There are three key reasons for this.

First, the dominating discourse in educational policy has been how to prepare young people for the economy of the future. This has led to disproportionately high investments by foundations and governments in the fields of STEM, with fewer resources available to invest in citizenship education. There is no doubt that in an innovation economy, STEM skills are necessary for a wide range of employment opportunities. But if these innovators and technical workers lack a deep connection to their communities, then there is little hope that they will be able to create solutions for the public good. It is true that voices have been quietly warning not to overlook the Civic Mission of schools (see the Carnegie Foundation reports of 2003 and 2011), with these voices growing in urgency and volume over the past few years—but they have not broken through to shape the overall priorities.

Second, citizenship education is hard. With citizenship education, schools must serve a double duty: they must provide a sense of comfort in a set of norms and a stable democratic system, and they must give students the tools they need to solve the challenges of today. Navigating this dual task is particularly hard in a world so marked by political polarization, racial divides, unvalidated news and high levels of migration. The additional stressors of climate change and pandemics are further exacerbating the divisions in society. Small wonder, then, that we struggle to define how to prepare young people for the future, when we are not aligned on the image of what future we should encourage our young people to see.

Third, the field is fragmented and siloed. With dozens of innovators offering complimentary programs, each with their own model of curriculum, instruction, and assessment, there has been no joint advocacy with a basis in analytics and research. Nor is there agreement on what it is that we should be teaching to begin with, let alone what we should assess. In the struggle for resources on local and state levels, educators who care about raising young people to be empowered citizens have not had data on which to base an argument for increasing their investment in teaching towards civic impact. Without this data, outdated teaching models of social studies and history that actively counteract the beneficial effects of citizenship education remain the norm, and funding for new priorities and innovation in the field of citizenship education across topic areas is in decline.
Fragmentation and reach

Citizenship education takes many different forms. First off, there is state-mandated civics learning, which is focused on ensuring a basic foundation of democratic literacy. Some form of civics or social studies education is now part of graduation requirements in every U.S. state, but the amount varies widely by state. Thirty-seven states require students to demonstrate proficiency through assessment in civics or social studies, and forty-eight states include civic learning as a strand in their standards (Education Commission of the States, 2016). In Australia, the national curriculum includes civics and citizenship as a subject under the Social Sciences and Humanities Learning Area, indicating that students, “will investigate political and legal systems, and explore the nature of citizenship, diversity and identity in contemporary society (Australian Curriculum, 2020).

Unfortunately, a state mandated implementation of “civics education” today sometimes means no more than cramming a civics exam or test into an existing system and expecting learning. This is not a solution. To simplify, the chain of events is often as follows: constituents note that young people are not civically informed or engaged. They pressure their politicians to act. Politicians then announce “we need civic education”, which results in a 50-question multiple choice test. Teachers are mandated to devote resources to ensuring that all kids pass this test, with no clear benefit to the student in terms of their civic capabilities. In many of the states that have recently implemented civics, this is the pattern that has been deployed, to the detriment of students and the frustration of teachers (NCSS, 2018).

Second, there is Service Learning. About 13 million students in the U.S. experience volunteerism in the form of Service Learning, a teaching and learning strategy that integrates meaningful community service with instruction and reflection, and which overlaps in important ways with the goals of citizenship education. It is offered in about half of public high schools, but not all students participate (Johanek & Puckett, 2004). A newer form of service learning or Civic Action learning has been gaining foothold with a rise in programs and interventions focused on encouraging the growth of young people as changemakers and entrepreneurs (e.g. Facing History's Choosing to Participate, Peace First, Harvard's Civic Action Project).

We have no data on how many students participate in such in-school programs nationally or internationally, and the research on its impact is still pending (D. Kidd, Personal Interview, 2020).

Finally, there are organizations focused on providing educational content through the lens of citizenship, either attached to core curriculum or as supplementary interventions. Most of the offerings of the leading providers in the U.S. (e.g. Facing History, World Savvy) are implemented piece-meal, leading to our conservative estimate of around 2% of young people in the US receiving comprehensive citizenship education (Team and BCG Research). But organizations can only track whether they have been used in schools. They have no way of knowing any details on the quality of implementation, nor can they routinely capture the results of the intervention. When programs can prove results, they can rapidly scale. As an example, High Resolves in Australia has reached an estimated 350K students in the Australian education system with a set of highly targeted interventions.

Thanks to a rigorous evaluation process baked into the delivery model, High Resolves was able to attract a grant from the Australian Government to reach 50% of Australian youth in the next four years. But overall, most civic education programs are teacher-selected supplemental materials with no connection to any systematic assessment model or reporting system, so tracking their impact and reach is hard.

The splintering of the citizenship education landscape has undoubtedly slowed down the collective reach of the field, as teachers struggle to select the best tools for a given goal, and can’t directly assess the effect on their students. High Resolves led an initiative during 2018 and 2019 to create a collaborative ecosystem across all civic learning providers that would allow teachers to create learning plans using the best available materials chosen for their ability to deliver against stated learning goals. The resulting platform, www.composereducation.org, was launched in April of 2020. Later on in this paper, we will describe the importance of this platform for the development of shared approach to assessment – at this point it is only relevant that there is light on the horizon in terms of the fragmented landscape, and that we have the possibility of scaling collaboratively as a field.
The Challenge with Measuring “Soft” Skills
When Joe Nye coined the term “Soft Power” in the late 1980s, he reframed how we see America’s influence in the world by showing that this influence derives primarily from things that are not easily quantifiable. Hard power, he argued, can be measured in men and missiles. Soft power, however, is what creates loyalty, aspiration and community—the things that effectively move goods and people. “What is soft power? It is the ability to get what you want through attraction rather than coercion or payments. It arises from the attractiveness of a country’s culture, political ideals, and policies.” (Nye, 2004)

But measuring soft power is hard, and it is most easily done through the lens of hindsight and history. Similarly, soft skills are what truly enable people to engage as citizens in a community with intent and direction, and these skills, like elements of soft power, are hard to pin down and measure. These are, however, the skills at the core of existing frameworks of civic education. Many states have been developing “profiles of a learner”. Around 30 U.S. states have adopted the “C3 framework” for social studies (College, Career and Civic Life) which focuses on developing young peoples’ capacity for inquiry and action, and references the broad goal of preparing a student to be an active and responsible citizen who can “identify and analyze public problems; deliberate with other people about how to define and address issues; take constructive, collaborative action; reflect on their actions; create and sustain groups; and influence institutions both large and small.” (NCSS, 2013). All of these goals would fall into the category of “soft skills”. All of these goals fall into the category of things we have struggled to measure in the past.

The predominance of hard skills assessment
Since the inception of assessments in schools, their focus has been on measuring hard skills. Parents want to know whether their child is “on track”, and want to know whether they are in a “good school”. Constituents want to know whether their taxpayer (or tuition) dollars are well spent. Schools want to know whether their teachers are performing. National policy makers want to know how their country compares with others, and whether we are achieving our promise to create opportunities for advancement for all. And so we created tests to assess children on the indisputably important core skills of reading, writing and arithmetic.

What any standardized test on core skills—which they are in-class achievement tests or aptitude tests—will find is that there are huge gaps between students, and that these gaps are not fair. Policy makers come into power vowing to close these gaps, and they have, in the past, used testing as one of their main tools to advance equality. The combination of the Bush administration’s policy of “No Child Left Behind” and the Obama administration’s focus on refining a “Common Core” intensified and refreshed the commitment in the U.S. to testing, in the hopes that the education system would then be able to diagnose weaknesses and advance solutions. Unfortunately, the tests failed to help achieve any advancement in equality of access, and relative mobility rates across income levels are no better in the US than they were 30 years ago (Chetty et al. 2017). Advances have recently been made in the field of assessing core skills thanks to innovators like Amplify, who have interwoven assessments with carefully designed curricula.
Ironically, standardized aptitude tests were created to level the playing field in education and to increase access to better education for all. As an example, the SATs were designed in 1926 to serve as pure intelligence tests so that young people without privilege could gain access to scholarships at elite institutions. In the century since, the SATs (and the ACTs that followed) came under intense scrutiny for being indicators chiefly of socio-economic status. As an industry evolved around providing coaching to guarantee high results in this privately administered, fee-based testing system, privilege was not dissolved, but cemented. Today, after almost 80 years in use as a central screening tool for colleges, these tests are on the chopping block in the debate over how to improve the fairness of college access (Sackett, 2019). As of May 2020, accelerated by the Varsity Blues scandal and the COVID pandemic, the largest state education system of the United States has stopped using the ACT/SATs in their admissions—others are expected to follow.

The move away from assessing young people on purely academic skills as measured by standardized tests has been quietly gaining momentum for years. Psychometricians and psychologists have been accumulating evidence for decades that personality traits and soft skills are better indicators of life success than standardized tests (Heckman and Kautz 2012). These insights have been corroborated and reinforced by the observations of educators. In the 2018 Gallup survey conducted in the U.S., only 9% of superintendents said that standardized test scores were “very important”. This same Gallup poll revealed that superintendents rank “soft skills” as more important than academics for preparing students for college” (Stringer, 2018). An August 2018 survey by Northwest Education Evaluation Association (NWEA) and Gallup found that over 80% of teachers, parents, superintendents, and principals in the U.S. believed that it is just as important to assess nonacademic skills as academic skills.

So if standardized tests have failed to create equal opportunity, and principals yearn for a way to capture non-academic skills, where do we stand on our ability to assess and communicate these skills?

Existing soft skills assessment tools

The answer is, simply, on shaky ground. Our assessment landscape in soft-skills is as fragmented as the landscape of education providers, despite efforts for over 25 years to advance the importance of these skills in schools. In 1994, a coalition of educators and academics formed a coalition to ensure that high-quality, evidence-based social and emotional learning (SEL) was established as an essential part of preschool through high school education (see www.casel.org/history). Although the emerging coalition (CASEL) works only with 20 districts directly as of today, it has triggered a number of subsequent initiatives and shifted the conversation around priorities in the field. (CASEL annual report, 2019).

These conversations recently turned to metrics beyond the scope of SEL and on a broader definition of citizenship education. In 2013, the Spencer Foundation funded a multi-year measuring civic learning initiative. In 2019, iCivics hosted a convening in January to discuss the state of assessment of civic learning and RAND released the results of a multi-year effort to build a repository of measures of hard-to-measure competencies. Nonetheless, the conclusion of the Whitepaper written in December of 2019 on the future of the civic learning ecosystem was that there is “neither consensus around a consistent list of outcomes people in the space are trying to achieve, nor a standard test to assess students’ civic knowledge, skills, or dispositions.” (Woodrow Wilson Foundation, 2019)

As of Spring of 2020, we have found no widely used, robust citizenship-focused assessment models in use in either Australia or the United States. Most of the more widely used measures are focused on social-emotional learning and are primarily self-reported, despite the fact that self-report measures are fraught with weaknesses and limitations. (e.g. reference bias. See Duckworth & Yeager, 2015).
There is, however, a rich array of existing assessment tools that we have reviewed in two ways: through a deeper dive with experts and practitioners on the most widely used assessments of citizenship related learning (see Figure 2.), as well as a preliminary inventory and analysis of 45+ citizenship-related assessments. We have included a sample of that inventory in Figure 2.

<table>
<thead>
<tr>
<th>Tool</th>
<th>Constructs Measured</th>
<th>Context</th>
<th>Cost</th>
<th>Reach</th>
<th>Company Structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>The Snapshot</td>
<td>Character Skills</td>
<td>Used by independent schools to measure character skills, to inform enrolment decisions. An optional add on to the SSAT (standardized test used for admission to private school; reading/verbal/math)</td>
<td>$35/student</td>
<td>16,000 students (2017-18)</td>
<td>Non profit</td>
</tr>
<tr>
<td>Panorama Education</td>
<td>SEL, climate and culture</td>
<td>Online platform for SEL assessment, surveys and other school data</td>
<td>$2500 annual &quot;set up&quot; fee + $1.50 per student ($2000 min./school)</td>
<td>7 million students, 8,500 schools</td>
<td>For profit</td>
</tr>
<tr>
<td>Hello Insight</td>
<td>SEL, Career Readiness, College Readiness</td>
<td>Company is Algorhythm an online platform for SEL and similar skill assessment</td>
<td>$5/student, $500 minimum</td>
<td>30,000</td>
<td>For profit</td>
</tr>
<tr>
<td>Mission Skills Assessment</td>
<td>Non-cognitive</td>
<td>Made by ETS; MSA is a web-based tool that uses multiple measures for each skill; institutional data only (no individual level reports)</td>
<td>No data</td>
<td>No data</td>
<td>Non profit</td>
</tr>
<tr>
<td>DESSA</td>
<td>SEL</td>
<td>Company is Aperture Education; standardized strengths based measure of SEL, web or paper based</td>
<td>$4/student</td>
<td>No data</td>
<td>For profit</td>
</tr>
<tr>
<td>XSEL</td>
<td>SEL</td>
<td>Online platform for assessing SEL and peer networks. 2 45-minute stand alone elementary SEL measures; Networker measures student classroom relationships</td>
<td>No data</td>
<td>50,000 youth as of 2017-18</td>
<td>For profit</td>
</tr>
<tr>
<td>PLTW (Project Lead the Way)</td>
<td>Career and College readiness</td>
<td>Simulations and situational based judgment assessment of skills in context</td>
<td>No data</td>
<td>400,000 youth planned 2018-19</td>
<td>Not for profit</td>
</tr>
</tbody>
</table>
Among the measures reviewed in Figure 2, Panorama Education's online platform has the largest reach: 8,500 schools – nearly 10% of all schools in the U.S, and 7M students. Educational Testing Services (ETS) has a web-based tool that uses multiple measurements to assess non-cognitive skills, called Mission Skills Assessment, and it has growing momentum among independent schools.

The Devereux Students Strengths Assessment (DESSA) by Aperture Education offers both a digital and paper-based version of their SEL assessment tools and has recently released a high school version. They also offer a teacher report on student behavior. Finally, xSEL Labs’ SELweb is one of the few to move beyond self-report to offer Situational Judgement Tests. They reached roughly 50,000 young people in 2017-18.

Despite their limitations, self-report surveys remain popular because they are relatively easy to implement. CORE districts, a coalition of the five largest districts in California, has conducted the largest systematic measurement of non-academic skills to date using Panorama’s surveys. Chief Strategy Officer Noah Bookman explained “Surveys are the best way to gather data from a breadth perspective, because they do not take too much time and you can roll in the socio-emotional assessment with other survey topics such as school climate.” (N. Bookman, Personal Interview, 2018) Panorama’s sample size and ability to provide comparison data sets gives more confidence to school customers. They’ve been able to scale thanks to simplicity of implementation and the ability to customize the assessment in a modular way—schools can pick and choose skills to assess, or even add their own modules on to the survey.

Despite what seems like a saturated marketplace of various providers of assessment tools, there is widespread interest among educators in better assessment options. Only 10% of teachers believed that the assessments used by their schools measured these skills “very well.” (Gallup and NWEA, 2019) Even for the most sophisticated current users of assessment, tools like DESSA “aren’t even close to what we need” (K. Dragon, Personal Interview, 2018).

Testing has become a formidable, multi-billion dollar industry, and investments have been expected to rise. Classroom assessments that monitor and provide feedback over the course of student learning, also understood as formative assessments, accounted for $1.6 billion of school budgets in 2016, compared to spending on end-of-course, summative testing (typically state-mandated standardized tests) at $1.3 billion. Spending on classroom-based formative assessments is expected to grow 30% over the next two years (Molnar, 2017; Skuse, 2016).

With decades of effort and a high level of interest from educators, how can it possibly be that we have not shifted the assessment landscape yet? Part of the answer surely lies in the strong interests in maintaining the status quo by companies that perform the standardized tests. But far more importantly, the answer lies in how the citizenship education field has or has not engaged actively in the process. No assessment product has yet been designed leveraging the insights of what works in the classroom; nor has any been designed with educators at its core. No one has yet been able to design a product for assessing citizenship that meets the needs of the stakeholders in a process that galvanizes the field of citizenship education to stake its claim as the banner carrier and guarantor of the stated mission of our public schools.
1.3. User Adoption: Characteristics of a solution that will be adopted

Interviews with educational leaders revealed that the existing tools for assessing soft skills were either too complicated, created room for teacher bias, too cumbersome, too rigid, too prescriptive—in short, not able to help teachers get a clear picture of their students’ learning needs. Noah Bookman of California’s CORE districts, who is also a member of the CASEL “Assessment Working Group” (AWG) underscored the need for objectivity, pleading for “direct measures of these [soft] skills vs. asking kids to report on these skills or teachers to report.” James Gallagher, VP of Education for Aspire Public Schools, underscored the urgency of developing “more advanced, performance-based assessment models” (J. Gallagher, Personal Interview, 2018).

Teachers longed for formative assessments that in and of themselves supported the students’ in their learning journey, as well as for tools that were objective and easy to use. There has been a near explosion in the growth of spending on formative assessments, which now make up almost ½ of the overall assessment budget in schools and is expected to grow 30% over the next two years (Molnar, 2017; Skuse, 2016).

In order to understand the needs of educators, High Resolves and Got History separately conducted a series of focus groups and interviews with principals and teachers to collect desired attributes of assessment methods. They all expressed dissatisfaction with the current tools, but were quickly able to articulate what they would like a new assessment model to provide, along the dimensions of why, what and how.

- Assessment must strengthen the WHY of their teaching. Teachers would like assessment not to be only a tool for reporting, but something that will support them in teaching towards their mission, i.e. it must be inspiring, guided by a vision of society thriving with the power of fully engaged citizens. That being said, they wanted the framework to be flexible so that it could be adapted in different contexts.
- Assessment must capture the benefit of WHAT is taught, i.e. using learning science and analytics to capture in real-time what students are learning and how their behavior is changing.
- Assessment must not limit HOW teachers engage with their students, i.e. it should be modular and flexible, with a very low barrier to being implemented in a classroom.

<table>
<thead>
<tr>
<th>WHY?</th>
<th>WHAT?</th>
<th>HOW?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Framework yes</td>
<td>Real-time formative feedback</td>
<td>Accessible and easy to use</td>
</tr>
<tr>
<td>but framework agnostic</td>
<td></td>
<td>Modular and flexible</td>
</tr>
<tr>
<td>Advances equity</td>
<td>True to the latest learning science</td>
<td>Collective intelligence over time</td>
</tr>
<tr>
<td>No cultural skew or bias</td>
<td>Measure knowledge and behaviour</td>
<td></td>
</tr>
</tbody>
</table>

No such assessment model currently exists, but there is no reason that we should not be able to design one. We have the learning science, and we have an emerging consensus on key soft skills that underlie citizenship education and civic wellbeing. Before we go into outlining the solution, however, we must first analyze what process would be needed to create such a solution in order for it to take hold across the system.
System Adoption: Characteristics of a process to get us there

In addition to not being designed to fit the overarching goals of educators in a classroom, assessment tools have been developed and deployed in a way that is not conducive to affecting the system. There are two key challenges with the current process: first, the development process of assessment tools is too complex, slow and expensive; and second, the process has not been approached in a collaborative manner to ensure that all stakeholders are engaged from the get-go.

Integrated multi-modal process

There is no doubt that the process of designing metrics and assessment tools for “soft skills” is conceptually and logistically challenging (Gensowski 2012). Following the established models of designing psychometric tests, the measurement of soft skills usually follows a very linear, engineered process. Typically, this process starts with the selection of a specific competence to be measured (such as independent thinking). Then, content experts conduct extensive literature reviews in order to disaggregate the competence into its constituent sub-constructs and related behaviors. Assessment experts then design a bespoke test to measure the cognitive understanding, affects and behaviors. These experts finally conduct validity tests to confirm that their assessment instrument is measuring what they think it is measuring. This instrument has to be properly calibrated and then, $5M to $10M later, is deployed into the field.

When these metrics are then deployed, they are often deployed with instruments that are lacking. Despite advances in the field of assessment, there is still a tendency to rely on a relatively narrow range of assessment formats such as multiple-choice questions. Many have challenged the validity of this form of testing, especially the implicit cultural biases they embody, but like the old definition of democracy, it is defended because it is better than alternatives. Research in behavioral economics is increasingly combining survey work with more interactive forms of response, especially for interventions aimed at measuring soft skills (World Development Report 2015, Banerjee et al 2019).

Instead of continuing to rely on mono-dimensional, expensive systems, we must design a multi-modal approach to assessment. Ideally, we would bake the process of collecting data into the students assignments and activities, so that we do not have to create a separate assessment mechanism. This would allow results to be combined from games-based assessment, situational judgment tests and various self-reporting mechanisms to help triangulate results for a more robust and accurate claim about how well a student has achieved a certain area of growth.

Stakeholder inclusion

The other key challenge with how assessment models have been developed is that the various stakeholders were not fully integrated into the process. The process of developing measurements on citizenship education has been driven largely by academics and education providers. The academics have usually been looking for ways to assess the impact of a specific intervention or seeking a new way to capture a specific learning goal (as described above). However admirable their work and necessary for the development of the field, these research efforts do not in and of themselves change how things work on the ground, as the adaptation of these measurements for the classroom and their distribution across schools go beyond the scope of research projects. Universities are not agencies of systems-change, and are not set up to ensure implementation.

Education providers, for their part, have been looking for ways to measure the impact of their work in order to thrive and survive in a competitive and resource-constrained field. They have—either alone, or with the support of researchers—designed assessment tools tailored to measure the learning effect of the intervention they offer. They need the data they collect to make the case for their worthiness of donor investment and to compete in the marketplace of educators. The purpose of their assessment work is to grow their organization, not to transform the civic culture of the educational system.

Education providers, too, are usually not set up to be eco-system changers, and will struggle to attract funding to do assessment work that is not related directly to the specific needs of their organization.
The result of this siloed marketplace is a wide array of bespoke approaches, which cannot be compared with one another, and that do not serve to inform an answer to the overarching questions of “how well does our school prepare citizens?” and “how well prepared is this student to thrive as a citizen?”

These are, however, the questions on educators’ and parents’ minds. We saw above from the user feedback that educators are dissatisfied with what they have. At the most recent large-scale convening on assessment in civic education the key conclusion—after extensive discussions with over 100 experts for multiple days—was that we need “further innovation in civic learning assessments [...] particularly assessments that are co-created with educators and/or developed with the needs and realities of schools in mind.” (CivXNow, 2019).

No amount of conferences will solve this problem, nor is it an academic research challenge. This problem can only be addressed through “ecosystem entrepreneurship” (Eyoel et al., 2020), with a committed team articulating a clear system-wide vision and providing direction and infrastructure, but including and magnifying the voices of the stakeholders involved. Parents, teachers, principals, social entrepreneurs, community institutions, policy makers and academics must all be heard in the process of vetting and designing the tools we will use to transform our system to better serve our youth.

Both Got History and High Resolves have a track-record of galvanizing stakeholders in this way towards a common goal: we outline below how we believe that this would work in a way that could provide a breakthrough solution.

### Conclusion: Lack of Assessment is Single Biggest Barrier to progress

The impact of this lack of both a system-wide model for assessing schools on their capacity to deliver citizenship education, and of assessment tools for helping individual students develop along the dimensions of citizenship is immense.

On the one hand, without a comprehensive and innovative assessment approach, we can never make efficient investments in civic education. As of today, vast amounts of money are being spent on schools trying to achieve civic learning without anyone knowing what the effect and relative benefit of each intervention really is. If we broaden our lens to include the related domain of social-emotional learning and for schools who do offer social-emotional learning helping students to become “more engaged citizens” was a cited as a reason for promoting it, then schools are spending between $21-$47 billion per year, which includes an estimated $640M on products and programs (Krachman & Larocca, 2017). In a world of scarce resources, spending this kind of money without knowing what works and why is unconscionable.

On the other, without a comprehensive assessment approach, organizations in the citizenship education field cannot argue to be deserving of more resources and attention. Without it, we cannot reverse the trend of declining funding for social studies and make the case for prioritizing investment in citizenship education over, say, textbook purchases, or more STEM. Without it, we cannot shift parents’ mindsets away from pressuring their children to succeed in a system of tests that are not only meaningless to their future, but detrimental to their wellbeing. Without it, we cannot shift the culture of our schools away from being competitive knowledge factories to being incubators of young citizens.

As such, the lack of an assessment model that advances and guarantees high quality citizenship education is the single biggest barrier to creating the opportunity for all young people to succeed and thrive. We know from other fields that launching concrete measurement efforts can accelerate collaboration and progress (Raine, 2006). The matter is of utmost urgency in the field of education. Every year that passes without an assessment model geared towards supporting citizenship education, yet another cohort of young people will experience learning that usually does not serve the stated purpose of the educational system. It is time for a breakthrough in assessment, both in what the product looks like and in how it is developed and deployed. We owe it to the young people in schools today.
2 The Big Breakthrough

We believe that this breakthrough in assessment is at last within our grasp. The approach we propose is to reimagine the entire process of assessing citizenship skills, democratizing the development of classroom assessments and putting their implementation fully in service of educators and students. From our research, we know that educators want and need assessment models that accurately and quickly capture and assess learning, and that are flexible and accessible in a way that enhances instead of limits teacher engagement in their classroom. Over the past ten years, High Resolves has experimented with different approaches, and in doing so has drawn valuable conclusions on what works (and what doesn’t) to assess citizenship and other soft skills (see box).

Impact Measurement Journey

This new approach represents the culmination in a journey to assess the impact and improve the quality of our own programming, building on previous lines of validation we have and continue to use.

Line A. Market Validation

Demand-side validation. The best evidence we have that our program has impact is that schools are willing to pay for it. Schools in all geographies are demonstrating a willingness to pay. We know that schools with small, limited or scarce resources wouldn’t pay if it wasn’t having impact.

Net promoter score. We measure participant engagement in our activities using the popular net promoter score. In 2020, we have averaged a net promoter score of 43 across all our deliveries, which is very strong. We consider Apple and Lexus which are around 60 as benchmarks.

Awards and recognition. We received the Patron’s Prize at Good Design Australia, which goes to excellence in user experience design. We won the Aspen Institute's McNulty Prize as well as the Schwab Foundation award for social entrepreneur of the year. These prestigious awards are external validations of what we do.

Line B. Market Feedback & Learning

Teacher feedback. We make a massive effort to measure how others see the program and how they would enhance it. Each year, we collect feedback from teachers and incorporate those insights into our curriculum.

Student surveys. We also engage in measurement of students through survey questionnaires: 95% of participants rate our program as highly engaging; 91% achieve the targeted cognitive outcomes; and 87% achieve the targeted affective and behavioral outcomes.

Student voice. We analyze artifacts from student workshops such as the I Resolve To cards that students fill out at the end of any delivery. These provide a strong indication of willingness to act.

Staff feedback. Finally, our staff, who do all the delivery engage in monthly “glow and grow” review sessions and their feedback, which is often the harshest, is taken into account in enhancing the curriculum.

Line C. Market-based Measurement

Software release schedule. Our pace of curriculum enhancement is based on the model of the software industry. We have a major release every year and upgrade the curriculum a great deal. This rapid iteration based on market feedback has been the key to our impact enhancement.

Independent third-party reviews. We have done independent third-party reviews of the program, which basically have established the effectiveness of the program. For example, ARC (Assessment Research Centre), a specialist unit at the University of Melbourne did such a report for us a few years ago and we could make that available to you.

RCT in Australia. We have proposed to conduct a more formal RCT evaluation as part of our next grant from the Australian government.
The result of our research and practical experience in the field is an assessment approach with four distinctive qualities:

1. It is anchored in a comprehensive citizenship education framework while remaining framework agnostic;
2. It shifts the focus away from the assessment of individual interventions to the continuous process of assessment and feedback of student growth at the classroom level;
3. It rejects the dichotomy of assessment and curriculum though the use of embedded micro-assessments;
4. It reverses the flow of how we build assessment, and collect data in an aggregate form.

We propose bringing these insights into a building an assessment system that achieves scale by operating across multiple, interconnected platforms as it proves itself and grows.

A Universal Framework but Alternatives Embraced

The Citizenship Quotient (CQ) Framework

The first key design element of an assessment approach must be a framework to conceptualize that we are talking about developing young people in a holistic way and to visualize the core learning outcomes and values to which we are committed. There are several frameworks circulating in the space of citizenship education with much overlap. We realize that a broad assessment project will require us to be adaptable across all of these frameworks, but we will walk through one framework for purposes of illustrating how we would manage to build a flexible system.

Since its inception, High Resolves has rooted its educational programs in a civic learning framework, which we developed as a means to reflect the core competencies of citizenship and ensure that we were orienting our learning modules towards delivering against clearly defined and measurable goals. We define citizenship as acting in the long-term collective interest: people must learn to think, feel, speak, and act like citizens and they must experience and drive towards both personal and societal transformation. Taken together, these two axes define the spectrum of core citizenship competencies and present eight distinct competencies which are foundational to citizenship.

We call this array, the Citizenship Quotient (CQ). Similar to the original IQ for intelligence, and EQ for emotional intelligence, CQ is a way of measuring the set of knowledge, skills, attitudes, values, and behaviors that enable people to work towards the long-term collective interest.

8 citizenship competencies \times 3 spheres of life = 24 expressions of citizenship competence
At the micro level, CQ provides students and schools with a visual map of their citizenship strengths and areas for growth. This enables principals to identify where gaps in citizenship competency lie in the school, identify learning interventions to address those gaps, and then measure the impact of chosen interventions.

At the macro level, CQ serves as a standard language across systems to determine the degree to which students are learning citizenship competencies as well as the effectiveness of different types of those interventions. By aggregating and leveraging data on CQ from schools across the world, we are able to identify effective learning interventions and gain deeper insight into what combinations of programs work. With this information, school systems will be able to redirect precious investment dollars to the highest-impact interventions.

Our framework is an expandable and adaptable one, with the key goal of putting civic growth and citizenship education into the front and center of any educational system. We have used the CQ Framework to help schools review how well they convey the skills and mindsets that cannot be conveyed just in a specific civics unit or in a history course, but in experiences and learning modules across all disciplines throughout a student's school career.

Framework Agnostic

However, we realize that there are a number of different frameworks in the citizenship education space, and it is not our goal to superimpose one framework on anyone. Educators all have preferences for particular frameworks or competencies when setting learning goals in their respective classrooms, schools, and institutions. These preferences are often expressed in the differences between frameworks, and even in the meanings of terms employed in them. For example, two schools interested in assessing identity might have related definitions of the competency that differ in detail, with one school opting to include a focus on collective identity at the global level, while the other focuses on the identities of the individual student.

Our approach is flexible and agnostic with regards to the framework educators used. On its face, this might appear to be a daunting task. Frameworks produced by groups in the citizenship education sector can vary in significant ways even while using similar terminology. By deconstructing the most common competencies into sub-constructs, the integral ‘building blocks’ of competencies, an educator could build their own frameworks to suit their particular definition of a competency, while another might choose a similar set, with some sub-constructs excluded.

There is already movement within the soft-skills sector to identify overlaps at the competency level. The Taxonomy Project at EASEL lab, an initiative led by Dr. Stephanie Jones at the Harvard Graduate School of Education, has mapped 40 frameworks that focus on social-emotional learning and soft skills against one another to identify overlap and alignment between them. A 2017 study by the American Institute of Research has also mapped all existing frameworks, identifying an urgent need to connect frameworks to practices and measurement (Berg, J. et al., 2017).

This movement gives us confidence to move forward with our framework-agnostic approach. What this would entail is identifying overlaps between some of the most popular frameworks at the sub-construct level and breaking down each competency within a framework into its relevant ‘building blocks’. Once a sufficient number of the highest-frequency sub-constructs have been identified, we would be able to select the set of sub-constructs that best fit the definition of a particular competency within a given framework.

The result is a model that accommodates many different frameworks and even the option for an educator to create their own, allowing for greater flexibility and specificity in what educators choose to assess.
Shifting the Focus from Single Interventions to Strings

Focus on Single Interventions is Not Enough
As described above, with no unified directive or mandate governing the delivery of citizenship education, educators have relied on a fragmented and dispersed approach that relies heavily on the expertise of third-party education providers. Unsurprisingly, when assessments are deployed to assess student progress on citizenship skills, they are frequently confined to the evaluation of individual interventions deployed by the provider involved. This approach is understandable. Education organizations are accountable for demonstrating impact, both to beneficiaries and funders, and therefore must demonstrate the impact of their programming.

However, this fragmented approach offers an incomplete picture of student progress beyond the scope of the single intervention, and creates a burden on educators to attempt to create and assess a unified program that draws on programming from several education providers. A new approach to assessment should shift focus away from single interventions, which by design are intended for use by those who designed the intervention, and to the continuous process of assessment and feedback that is a staple of responsive teaching and growth-oriented learning.

Strings-Based Approach to Assessment
Our approach relies on the design and delivery of customized sequences of learning experiences, which we refer to as “strings.” Each string consists of different kinds or types of experiences, called elements, which incorporate insights from the field of learning science on best practice in curriculum design. These types include: Peak Experiences, Repeated Practices, and Real World Applications (see box). When educators combine elements from different providers (each with its own unique color) into a coherent string, they will be more effective in helping students build mastery.

Mastery Formula Element Types

**Peak Experiences** are immersive or affective ‘ah-ha!’ moments that introduce new knowledge and can provide motivation or a sense of urgency around the pursuit of a new skill or learning objective. These can include workshops and simulations, museum visits, and powerful films. Peak Experience elements are represented as triangles.

**Repeated Practice** are opportunities to practice a new skill or competency and receive feedback. Practice often begins with the fundamentals, but increases in difficulty and complexity over time to help transfer knowledge and behavior from working to long-term memory. These can include activities, lessons, case studies, and exercises. Repeated Practice elements are represented as squares.

**Real World Applications** are extensions of effective practice, but which move beyond the confines of the classroom into the ‘real world’, providing students with opportunities and confidence to apply their skills to new challenges in ways that are both meaningful and practical. These can include community engagement projects, awareness campaigns, and service learning. Real World Application elements are represented as circles.
String Example

We call this approach the Mastery Formula. Informed by insights from the field of learning science, the framework encourages educators to design for a particular set of outcomes, scaffolding skills and competencies towards explicit objectives and using the different types of elements and experiences to reinforce and capitalize on learning opportunities. At High Resolves, we work with our school partners to develop these customized strings based on the unique citizenship educational goals of their students, allowing for a wide array of different interventions. These strings incorporate not just materials produced by High Resolves. Rather than invest time and resources in the development of materials, we promote the use of resources created by other education providers whose work we value, and whose materials we have no wish to duplicate. The result is the creation of “hybrid” strings, with each provider represented by its own unique color, that draw on a variety of experiences, but which are unified in purpose and objective.

Our approach to assessment aims to assess student learning at the string level. By keeping the focus of assessment on strings rather than the individual offerings within them, our assessment tool will be able to support educators in their mission to assess and offer real-time, useful feedback to their students.

Rejecting the Dichotomy of Curriculum and Assessment

Bridging the divide between the lesson and the test

The orthodox approach to assessment assumes a hard separation between the lesson and the test. This dichotomy between curriculum and assessment is a false one. Why does the test have to follow the lesson? Why can’t the assessment be fully embedded into the lesson itself? Why can’t we use students’ participation and behavior in a learning experience to support claims about students’ competencies?

In order to do so, we must first loosen the grip that multiple-choice questions still hold on people’s imagination and practice in assessment. The validity of this form of testing has long been put in question, particularly the fragmentation of knowledge it encourages and the implicit cultural biases it embodies, (Elford, 2002; Busteed, 2016) but like the old adage goes about democracy, it is defended because it is seen to be better than alternatives. It is clearly the easiest format of assessment to deploy to get values for comparison across all learners and school types.

At its core, however, it goes against the now mainstream concept of “Multiple Intelligences” as defined by Howard Gardner, and norms our teaching towards forms of learning that do not engage, develop and empower students. Different children learn in different ways, which means that we need multiple inputs to form a complete picture of how their competencies have advanced. Research has established the value of using multiple modes of assessment to enhance the robustness of the result (Mattern, 2012). Indeed, the best degree of fidelity can be achieved by triangulating three different kinds of information:

We have been exploring how we could leverage our games, activities, and simulations to collect data on how well students are learning and demonstrating certain behaviors (see box). What we have found is that we can integrate observation and evaluation of growth into a wide variety of activities, and can collect data without a summative assessment.
### Leveraging Lessons for Data Collection

In the Partial Truths activity, students take on the role of two companies attempting to sell an energy drink using a set of facts that include positive and negative information about each company. In our experience, students will tend to use all the facts that benefit their brand, or the fact that are negative for the opposing brand, while ignoring the data that is less beneficial. Discussion reveals to students the way partial or selective presentations of the truth can be used to manipulate others.

As an assessment, the activity provides multiple opportunities to observe and collect data on student behavior.

For example, the facilitator can observe students’ participation, decision-making, as well as articulation of key learning outcomes.

These student behaviors can then be mapped against skills associated with the competencies of a particular framework for assessment.

### Breaking the False Dichotomy: Curriculum as Assessment

This leads to the third distinctive quality of our approach, which relates to how we capture information to assess the impact of a learning module or intervention on the development of CQ. Rather than accept the traditional view of assessment, we have piloted embedded micro-assessments (EMAs) as the answer to competency instrumentation and assessment design.

EMAs are like pulse checks—less rigorous, but seamlessly woven into and tailored to specific curriculum (see image). This approach minimizes any specific “assessment events” so that assessment and learning become a single experience. These pulse checks are bite-sized, modular, and able to be done in 5-15 minutes. The approach allows for rapid, short cycles of feedback.

---

For example, in our Social Progress workshop, we run the CO₂ Game-- a complex, immersive simulation around global climate change negotiations. Students play Ambassadors from a few dozen countries in a multi-party prisoners’ dilemma, working to reduce global emissions while reducing the cost in carbon credits to the countries they represent. As an EMA, this piece of curriculum would act as both a learning experience and assessment opportunity. We could track not only the students’ formal emission reduction decisions but also their game strategies and many aspects of their conduct through the game. This data could be matched up against many of the required data points for formative assessments of citizenship competencies.
At scale, this approach would transform the practice of assessment. The sum of EMAs over the course of a year would be interconnected, so that each formative assessment builds on and connects to others, to enable “cumulative validity” meaning that, by regularly assessing learning over time, by the end of the year, the formative assessments as a whole cumulatively enables summative understanding of student progress towards citizenship education learning goals.

To an excellent educator, this process of continuously collecting information about a student’s growth in different ways is second nature, and in small progressive schools where close attention can be paid to each student and standardized assessment is eschewed, this is how young people are evaluated in their development as learners. In order to make this a broad and accepted norm, we must create new ways to capture this information across larger groups of students and across many different types of learning intervention. This will require collecting data on the impact of learning experiences on students at scale, which leads to the next innovation in our assessment approach: reversing the flow of assessment.

**Reversing the Flow of Assessment**

**The Typical Flow of Assessment Design**

When designing an assessment, psychometric experts typically start by choosing a specific competency to be measured. Then, content experts conduct extensive literature reviews in order to disaggregate the competency into its constituent sub-constructs and related behaviors. A routine part of assessment development is a specification of what the assessment is intended to measure. At a minimum, this should include: (a) operational definition of the competencies to be measured by the test; (b) a listing of its sub-constructs; (c) learning progressions, and (d) a listing of possible observable behaviors.

As described briefly above, assessment experts then design a bespoke test to measure the cognitive understanding, affects and behaviors. These experts finally conduct validity tests to confirm that their assessment instrument is measuring what they think it is measuring. This instrument has to be properly calibrated and then is deployed into the field. This process is summarized in Figure 1.

![Figure 1. Traditional Assessment Process](image)

However, this very expensive undertaking is often poorly suited to educators and students. The development of these assessments take a long time to develop, are often difficult to use or deploy as a manner of routine in the classroom, and are frequently restricted in scope to single frameworks or competencies, which an educator may not use or define differently.
Reversing the Flow of Assessment Design

We propose a radical new approach that reverses this flow. Instead of creating activities to measure the sub-constructs of a competency, we propose to begin with a set of extant immersive activities and analyze which behaviors they display, collecting this information in a large database. In essence, we would be doing with educational experiences what sensors and the Internet of Things (IoT) has been doing to physical assets in other sectors. We could then match these behaviors with pre-identified behaviors already linked to sub-constructs to allow assessments to be created very quickly and cheaply (see Figure 2).

![Step 1 Diagram]

In order to demonstrate how this framework would be applied in practice, we operationalized the High Resolves CQ competency around Collective Identity. This work was achieved through a literature review which surveyed over 130 sources.
We began by identifying the measurable components of the competencies that could potentially be observed in the in-game behavioral data, creating a concept map or a comprehensive representation of given competencies that consists of both high-level and low-level concepts and the relationships between those concepts. This map is shown below.

In this approach, we identified a set of eight sub-constructs that research suggests comprises collective identity (as defined by High Resolves), or the ability to appreciate our shared humanity as members of a single human race. These sub-constructs were grouped in three categories, which proceed from the cognitive to the behavioral and these sub-constructs build upon other sub-constructs both within and outside their category.

Then, manifesting behaviors were identified for each of the categories of sub-constructs. Manifesting behaviors for each category were likewise organized into the three different spheres of life employed by the CQ framework, demonstrating how a cluster of sub-constructs might manifest as observable behaviors in the different areas of life where collective identity skills are employed.

We envision that rigorous work will need to be done to ensure meeting the psychometric standards of the assessment. However, we suggest that a first step this effort is to provide conceptual clarity—in other words, an operational definition that provides sufficiently concrete terms to help inform measurement efforts.

Up to this point, our approach is quite traditional. We take a complex construct and break it down into sub-constructs that you could measure with behavioral data. The novelty of our approach lies in the next step, in which we reverse the process of gathering data. Instead of creating new activities that would yield opportunities to observe student behavior, we propose identifying a set of immersive activities for which the required behaviors are already present, or would be with minor modifications. In this example, we would then connect these concrete activities with the Collective Identity competency which has already been deconstructed into relevant behaviors. Through a mix of EMAs in the classroom, situational judgment tests and various self-reporting mechanisms, an assessor will be able to form an increasingly complete assessment of students’ sense of collective Identity.

We believe 5-8 immersive activities would be sufficient to assess dozens of citizenship competencies and the entire CQ model. That would mean a school could incorporate a few of these immersive learning experiences into its curricula, and get formative assessment feedback on dozens of citizenship competencies without the need for any stand-alone testing. Market research and our focus groups tell us this is exactly the type of assessment and measurement solution that principals and school leaders are looking for.
Conclusion: Disruptive Model for Assessment

The above example represents the initial steps taken to operationalize a single competency from the High Resolves CQ Framework and a set of EMAs. However, the potential applications of this approach transcend any single organization's framework.

The four distinctive qualities described above comprise a new assessment approach that could be applied to any number of citizenship education frameworks from providers, schools, and policymakers from around the world. By instrumenting up both frameworks and a set of EMAs using a common list of recognized sub-constructs, we can create a menu of EMAs that can be curated to create bespoke assessment plans for whichever framework an educator, school, or organization might prefer.

This approach has the potential to significantly disrupt the way we assess citizenship education and meets all the requirements of an effective solution. It is flexible enough to accommodate the variety of frameworks found in the sector; it obviates the need to create new assessments from scratch for each new framework; and—perhaps most importantly—it provides prompt feedback to educators on student growth.

Sceptics might counter at this point that it is absurd to try to create a highly scientific assessment model for something as broad and individual as citizenship qualities. To that one must answer that we don't have the option in a state-run educational system to not measure the kind of learning that matters most. What we are measuring now is making young people stressed, not helping them succeed as individuals and not helping to close the blatant gaps in opportunity. So it is not a luxury, but a moral imperative to design a scientifically rigorous and robust measurement model on the citizenship qualities that matter most to individual, societal and planetary wellbeing.
Bringing the Breakthrough to Life

Multiple, Reinforcing Platforms as a Technological Solution

Earlier, we described the nine characteristics of a breakthrough solution that underpin the approach detailed above. A breakthrough solution for the citizenship education sector should be modular and flexible, customizable to the needs and preferences of classrooms and educators, accessible and easy to use, with the capacity to process large amounts of data at scale.

Transforming the approach we’ve described into a practical and accessible tool will require a suite of interconnected technologies that will seamlessly allow educators to 1) discover high quality citizenship education content and create customized learning sequences for the classroom; 2) find or create EMAs based on the desired outcomes of strings; and 3) collect data during instruction for fast analysis and prompt feedback.

The development of a tool that satisfies these requirements is no small task, requiring a digital ecosystem that touches on every aspect of the learning experience. However, over the last three years, High Resolves has built or partnered with existing digital platforms to support educators in the discovery, design, and delivery of citizenship education curricula.

We are now proposing a new assessment engine, to work in concert with our existing and emerging platforms, Composer and OpenLearning, to bring this breakthrough in citizenship education assessment to life.

Composer: The Learning Experience Platform (LXP)

Composer is a digital platform and collaborative ecosystem that makes it easy for educators to discover, access, and coordinate top citizenship education resources from around the world. Using the latest insights from learning science to aggregate and organize content, Composer offers users a robust library of resources from several dozen leading education organizations and a suite of curriculum design tools to organize that content in comprehensive learning sequences.

The vision of Composer is to support educators as they guide students through their growth as citizens, individuals who act for the wellbeing of their communities and the planet. To support educators and students in the development of citizenship competencies, Composer has aggregated content covering a wide variety of themes, broadly organized into four overlapping domains: civic learning, social justice, social and emotional learning, and global competence. The result is a robust digital library covering a wide variety of soft-skills frequently associated with citizenship education.

The platform’s combination of easy to access resources with its modular approach to curriculum design allows for customization of content to a particular classroom context, supporting educators in the integration of citizenship education into school programming. Composer has adopted the Mastery Formula and strings-based approach described above, aggregating content from a wide variety of providers, represented by different colors, and categorized into the different types of elements: peak experiences, repeated practices, and real world applications (see Figure 1). This approach offers a common lexicon for content uploaded by different providers, allowing users to search for, organize, and create strings, combining content from many different providers.
The platform is the product of a collaboration between five founding provider partners, recognized for excellence in citizenship education: High Resolves, Facing History and Ourselves, iCivics, Generation Citizen, and Peace First and funded by the generous support of the Chan Zuckerberg Initiative and Omidyar Network. Composer was launched to the public in April 2020 and continues to be managed by High Resolves on behalf of the ecosystem until such time as it is ready to become its own independent not-for-profit entity. The ecosystem has grown to include over thirty education providers and approximately 1000 unique learning experiences.

In an integrated platforms approach, Composer acts as a learning experience platform (LXP), where users discover, search for, and create sequences of citizenship education that can be assessed through the use of EMAs. Composer offers several advantages to this end:

**Robust Digital Library.** Composer has curated a library of content from leading organizations in the citizenship education sector, offering a ready list of prospective activities that might be mapped using the embedded micro-assessment approach and included in educator curriculum plans.

**Ease of Integration into Existing Architecture.** Composer has already built a powerful curriculum design tool that allows educators to integrate materials from a wide variety of organizations into their curricula. Educators would be able to easily identify and integrate embedded micro-assessments into a scope and sequence.

**Market Access.** Composer is built for scale, creating access for educators around the world with access to high-quality citizenship education materials. Composer would make high-quality assessment available to educators at low cost.

### Table: Composer Elements

<table>
<thead>
<tr>
<th>Elements</th>
<th>Providers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Isle of Dogs</td>
<td>High Resolves</td>
</tr>
<tr>
<td>Identity and Purpose</td>
<td>High Resolves</td>
</tr>
<tr>
<td>Purpose Compass</td>
<td>Project Wayfinder</td>
</tr>
<tr>
<td>Active Listening</td>
<td>High Resolves</td>
</tr>
</tbody>
</table>

**Figure 1. Elements on Composer**

<table>
<thead>
<tr>
<th>Title</th>
<th>Provider</th>
</tr>
</thead>
<tbody>
<tr>
<td>Isle of Dogs</td>
<td>High Resolves</td>
</tr>
<tr>
<td>Identity and Purpose</td>
<td>High Resolves</td>
</tr>
<tr>
<td>Purpose Compass</td>
<td>Project Wayfinder</td>
</tr>
<tr>
<td>Active Listening</td>
<td>High Resolves</td>
</tr>
</tbody>
</table>

*The Future of Assessment White Paper 2 | 24*
OpenLearning: The Learning Management System (LMS)

OpenLearning is a digital platform revolutionizing online learning to be a social and engaging experience. Founded by Adam Brimo, Richard Buckland and David Collien in 2012, the Sydney-based, ASX-listed, software-as-a-service company that provides a scalable online learning platform and learning design services to education providers; and a global marketplace of world-class micro-credentials and online degrees for learners. It is what educators call a Learning Management System (LMS).

OpenLearning offers a suite of tools to build online courses that require learners to take an active role in the learning experience, rather than the tedious cycle of information consumption and memorization that plagues most e-learning courses. The site uses widgets and features inspired by social media to create communities of learners, who explore topics, skills and concepts through active decision-making and collaboration.

Built on proven learning sciences research and a social constructivist learning philosophy, OpenLearning goes beyond traditional instructivist approaches and static learning management systems to deliver authentic, active and connected learning experiences. With over 62 universities, corporations and government agencies delivering over 7,900 courses to more than 1.74 million learners worldwide, OpenLearning enables education providers to expand into new markets, improve outcomes and diversify their revenue.

OpenLearning’s vision is to improve access to quality education, promote lifelong learning, and future-proof the workforce by enabling education providers to design, deliver and sell transformative courses and degrees worldwide. In the Spring of 2020, High Resolves partnered with OpenLearning to translate our signature peak experiences, traditionally delivered in-person, into engaging online workshops. This collaboration also presents an opportunity for High Resolves to bring the embedded-micro assessment approach to life.

The platform’s capacity to facilitate and capture student decision-making through online games, simulations, and collaborative work makes it an ideal candidate for a Learning Management System (LMS) that will allow educators to capture data on student behavior as a formative assessment.

For example, a teacher might choose to include in a string the Just Society workshop by High Resolves, facilitated through the OpenLearning platform. During the workshop, students take part in the Draw The Line simulation which explores issues of inequality and structural inequity on the planet. This simulation, instrumented according to the EMA approach, would collect data on student behavior: collaboration with peers in virtual chat rooms, changes in student prosocial behaviors over the course of the simulation, etc. Taken together with other data sources, the educator can use the simulation to form a picture of student progress towards skills associated with the activity during the validation process, such as the ability to work together with their peers towards the common good.

Obviously, not all Composer elements will be delivered via the OpenLearning platform. Nor will all of OpenLearning’s content be related to citizenship. In this context, however, we imagine using the LMS interface of OpenLearning to test the validity of our assessment tool and the formats in which experiential citizenship related learning can occur in an online environment. This information will be relevant to all online learning providers. By adopting similar backend coding, other providers of online learning could match sub-constructs to observable behaviors, and could feed the results into an assessment engine.

New Assessment Engine

The Just Society workshop is one example of an entire library of EMAs that could be facilitated through platforms like OpenLearning, and included by educators in strings through Composer. However, an additional platform is still needed. This additional platform would allow users to curate assessment plans that align with their chosen learning outcomes, selecting a number of EMAs to be embedded into strings as formative assessment for that purpose. The platform would also require the capacity to process data assessed during EMAs, offering near-immediate feedback on class progress in targeted skill areas and allowing educators to track student and cohort progress over time.
This new engine would integrate the novel competency mapping method of the EMA into the site’s library or menu of assessments. We would start with a selection of the most popular frameworks in citizenship education mapped against all content on the site at the sub-construct level, providing educators the opportunity to search for content using their preferred framework. At the same time, we would instrument a selection of games, activities, simulations, etc. as EMAs, allowing the site to collect data on the manifesting behaviors linked to one or multiple of the sub-constructs covered by the frameworks offered on the site. This approach would create a ready menu of relevant assessment activities, curated to meet the requirements of the competencies addressed by the educator’s preferred framework.

In this example, educators would begin by selecting a preferred framework and indicate the targeted competencies that will serve as the outcomes for their strings. Search filters would allow educators to access content relevant to their selected competencies and real-time feedback on curriculum design, informed by insights from learning science, would equip educators to design robust strings that successfully address the chosen competencies.

As educators build their strings on Composer, they would have the option to choose from the games, activities, and simulations that have been pre-instrumented as embedded micro-assessments on the assessment engine. If we extend the metaphor of the string, treating the string itself as the treble clef for a musical piece, then the parallel assessment plan would act as a support, keeping time with the melody of the string’s scope and sequence (see Figure 2).

By incorporating several of these pre-instrumented activities as lessons at strategic moments, educators would then be able to embed these activities as formative assessments into the scope and sequence of their strings. For example, a student participating in our Just Society workshop can complete all or portions of the workshop’s activities on OpenLearning, which have been instrumented for data collection in advance as EMAs. In one activity, students form family groups, completing challenges and puzzles that will allow them to accumulate points which they can then use to purchase basic necessities like shelter, food, and water. There are four versions of each necessity, and the more points students have, the higher-quality version of a resource they can purchase. However, what students do not realize is that the cost for these items are different for each family, reflecting the reality that for many people around the world, effort does not guarantee access to the same resources, opportunities, and benefits enjoyed by others. After completing the activity, students are then asked to ‘draw the line’, indicating which of the four versions every human should have.
As students participate in the activity on OpenLearning, the platform has multiple opportunities to collect data, such as the types of resources students choose, where students choose to draw the line after discovering the disparities in costs, as well as written student responses on their choices and reflections. The information from this EMA delivered via OpenLearning would then be fed back into the assessment engine, where assessment data would be analyzed and then reported back to the educator as near real-time feedback, creating a report on class progress towards the desired outcomes (see Figure 3).

The above represents our proposed application of the EMA framework as a technological solution. These three platforms working in tandem would enable educators to coordinate their curriculum design process with their assessment plan and content delivery, tightening the loop between planning, delivery, feedback, and iteration. The result is an educational experience more purposeful in skill development, and more attuned to student progress. In the following section we offer a proposal for a pilot of the application, and our vision for scaling the EMA approach.
Systemic change through ecosystem activation

This project of reimagining the future of assessment is no less than an effort at changing a system. In a recent article on the mechanisms of systems change, Kania, Kramer and Senge elegantly outline the necessity of addressing structural, relational and transformational change in order to “shift the conditions that hold a problem in place” (Kania, Kramer, Senge, 2018). With this in mind, we will address all three levels as we move to change the reality in classrooms as quickly as possible.

The tools and processes we describe above are located on the level of explicit, structural change. But we will also be addressing the relationships and power-dynamics that are responsible for the use of the tools that exist today. We will surface and outline the role of text-book companies and testing agencies in maintaining the status quo, and will analyze the power dynamics created by the economics in the citizenship education space. These are the challenges that are not discussed in academic conferences, but that are a key factor in preventing any new ideas from taking hold.

We will create opportunities—both with the newly discovered collaborative capacities online and in person—to weave a community of intent, building relationships and connections among the many providers and educators with a shared vision for a thriving democracy and wellbeing at an individual, societal and planetary level. We will do so by working with Composer and Got History’s international network of citizen education providers, to ensure that the models we are developing enhance and magnify their efficacy and impact.

We will also test our work in regional ecosystems, where museums, libraries, and parks play a supporting role in delivering civic education experiences. In Got History’s pilot Heartland Learning Collaborative we will test our emergent hypotheses in constant feedback loops with the stakeholders. One key goal of the project is to make it easier for educators to provide comprehensive citizenship learning: at every step will be collecting their input to ensure that the result delivers just that.
The other key goal is, of course, to ensure that all young people receive the learning experiences they need to develop full “participatory readiness”. They, too, are key stakeholders of the project. Therefore, a central feature of our ecosystem building process will be the inclusion at every step of young people. Got History is the sponsor and partner of a newly formed National Youth Council on the Future of History Education which will bring a collective voice into the project. These young people, ages 16-19, are deeply invested in ensuring that the students that follow in their footsteps are not limited by broken assessment models and curricula in their sense of power, belonging, connection and community. There is no one better than young people to help shape the kind of assessment tools that will allow them to flourish as citizens of the world.

Most importantly, we will build on and reinforce a shared mental model of citizenship education that focuses on supporting students in their growth as citizens of a community. This model is not new, and it is woven through the language of school districts across the globe. Nonetheless, it has failed to become the lived experience of students--to a large extent because of existing assessment models. Our assessment approach has focused every conversation regarding education on “closing performance gaps” and “job readiness” as a path to more equity and opportunity. The failure of that approach is on plain view, in protests that are covering the face of the globe in wake of the murder of George Floyd. People are on the streets demanding a new mental model for education, focused not on enabling higher performance, but true equality, inclusion and empowerment for all young people in the community.

Alongside our commitment to connecting and engaging partners around a shared mental model, the key success factor will be our ability to quickly deliver and test a practical prototype of a better assessment model that brings our shared vision to life in a concrete form.

**Proof of Concept Project Plan**

Our initiative’s core project in year 1 will be executing a rapid prototype and testing plan for the concept described above so that we can engage the ecosystem not on lofty principles, but on concrete tools for use in the classroom. We propose beginning with the mapping of the remaining domains from the High Resolves CQ Framework against content on Composer, along with the creation of several embedded micro-assessments facilitated by partners and through Open Learning that will then be tested for rigor and incorporated into strings. The proposed prototype includes the following phases:

**Phase 1.** During the first phase, we will begin research of the remaining CQ domains, developing operational definitions of competencies, their sub-constructs, and observable behaviors. At this time, we will conduct outreach to communities and users to gauge interest in pilot participation and what subjects they would like to see addressed in a string for the pilot. Several strings will then be developed, drawing on content from multiple providers.

**Phase 2.** In the next phase, we will develop an evidence model and build prototypes for the embedded micro-assessments. The first prototype will comprise 5-8 sets of micro-assessments embedded into the string of learning experiences for the project. For each activity, we will observe student participation to determine what evidence can be collected from the activities and how they might be modified to optimize data collection. During this time, we will conduct formal enrollment into the project, schedule key pilot activities, and train project leaders.

**Phase 3.** In the final phase, we will begin delivery of the program in participating schools and other partners (e.g., museum visits) including online elements over OpenLearning portals dedicated to each school. These small-scale trials will allow us to collect data on the EMAs, the results of which will be reviewed with experts and stakeholders. We will also evaluate data on participants’ experiences and responses to the prototype and process.

This integrated approach which blends the curriculum and platforms that High Resolves has built with Got History’s “ground-game” to excite and coordinate implementation partners is the magic combination to take our work to the next level.
The purpose of the proposed pilot is to validate our hypotheses of the embedded micro-assessment approach and to demonstrate the efficacy and impact of mapping a competency, its sub-constructs and observable behaviors against activities selected to act as EMAs. Through this process, we hope to learn what potential adaptations might be needed to existing activities, and the methods for data collection, to ensure EMAs function as rigorous assessment tools. Learnings from this pilot will then allow us to revise, iterate, and design for scale.

A Vision for Scale

With a pilot demonstrating how to capture the impact of elements of citizenship education, we will be able to start answering some very important questions in the field today: what experiences and curricula lead to the development of participatory readiness? What are the kinds of interventions best suited to learning in different environments? Who is receiving the kind of learning opportunities that they need, and who needs to be getting more? What are the most cost-effective interventions, and how can we ensure that they become the new norm?

Armed with answers to these questions—and only then!—the citizenship education community, including teachers, civic education providers, museums and parents—will be able to engage in a robust conversation with policymakers and budgetary decision makers on allocation of resources in education. Only then will we be able to start conversations about redesigning curricular standards so that the needs of students can truly be met.

We envision the process of scaling the pilot as one driven not by push, but by pull. If we are able to build an assessment platform that answers the unmet needs and performs the jobs to be done of the stakeholders involved, the platform will be eagerly adopted and spread by educators eager to highlight the impact of their teaching and schools eager to show the power of their educational offerings. While we intend to develop and pilot this assessment platform in connection with Composer, we envision the assessment model we develop as being platform agnostic, and creating a new standard. We hope to be a solution that leads the way rather than being the only solution.

In the end, the breakthroughs we propose in the assessment field will serve not only to accelerate the interventions and ideas of the leading citizenship education providers for higher impact. We will also be paving the way for those educators who have been teaching to these citizenship goals all along, in whatever format they have chosen, to become the dominant model of teaching. Having a framework agnostic, platform agnostic model that can connect abstractly defined citizenship learning goals with concrete interventions and proof of impact will allow the conversation in education to shift to where it started at the beginning of human formal education: towards equipping the next generation with the knowledge, tools, and dispositions needed to act for the wellbeing of themselves, their communities, and the planet.
Conclusion

Our vision obviously goes farther than the scope of this 2-3 year initiative, as we can actually see today in vivid colors the full impact that this project could have..

We see a future educational system transformed by the power of reimagined assessment. What if, in the “Best School Rankings” of 2030, we had a central rubric reflecting not how many students a school delivered into Ivy League Colleges, or how their graduates performed on average on the state's standardized tests, but how many students graduated proficient in citizenship? What if principals touted their schools' demonstrated capacity to deliver learning against key citizenship goals, with data and curricular choices to bolster their statement of mission and values?

What if we had a public conversation revolving not around “achievement gaps”, but about our shortcomings in ensuring the civic wellbeing of the young people in our care? What if journals were full of articles reviewing recent decisions by elected officials to invest significant amounts of money into citizenship education, providing teachers with the resources they need to select the tools and interventions that will best serve their students towards achieving citizenship-related goals?

Among those articles are also reports on reduced teacher-turnover and increased job satisfaction, as teachers revel in the newly felt freedom to focus on supporting the growth of each young person as a human being. Teachers delight in being able to show that the choices they made to select those learning units that fit the needs of their students and their ability to compile them in a sequence towards a clear learning goal made a real difference in the lives of the young people they teach.

From the chorus of youth voices magnified through instruments of civic engagement and empowerment we hear sighs of relief over not having to be on a hamster wheel towards a meaningless goal. They express excitement that their peers' marches and calls for change were not in vain, clearing the path for them to enjoy a school system that serves their need to grow up ready to engage as citizens for the wellbeing of all.

In the community of civic education providers, which has flourished through a system-wide validation of the best of their materials, we see a new level of differentiation and complimentary thinking, as providers hone their offerings to deliver those pieces of the learning mosaic at a very high level that have the most potential to impact specific groups of young people. They have been able to enhance and refine their offerings because of the years in which they deployed and tracked their impact across multiple learning environments.

This is a bright and inspiring future vision. Wouldn't that be the future that our children deserve? We invite everyone to be part of this conversation and process. We are committed to pursuing this vision with a clear product-focus, scientific rigor, a collaborative heart and open mind.


List of Individual Interviews – Market and Landscape Research

• Rosemary Benda, Social Studies Teacher, Oak Park High School, Detroit, Michigan
• Larry Berger, Chief Executive Officer, Amplify
• Noah Bookman, Chief Strategy Officer, CORE Districts, California
• Katie Buckley, Managing Director of Research, Transforming Education
• Jennifer Charlot, Partner, Transcend
• John Cronin, VP of Education Research, NWEA
• Kriste Dragon, CEO, Citizens of the World CMO
• David Feldman, Head of School, The Roeper School, Birmingham, Michigan
• Julia Freeland Fisher, Director of Education, Clayton Christensen Institute
• James Gallagher, Vice President of Education, Aspire Public Schools
• Laura Hamilton, Associate Director, RAND Education, RAND
• Peter Haun, English Teacher, Oak Park High School, Detroit, Michigan
• Erica Hodgin, Associate Director of Citizen Engagement Research Group, University of California at Riverside
• Sarah Elizabeth Ippel, Founder and Executive Director, Academy for Global Citizenship, Chicago, Illinois
• Kelly James, Partner, Education First
• Joseph Kahne, Faculty, School of Education & Director of Civic Engagement Research Group, University of California at Riverside
• Allison Kerr, Executive Director of Learning & Innovation, Citizens of the World Charter Schools
• Sara Bartolino Krachman, Co-Founder and Executive Director, Transforming Education
• John Rogers, Faculty Director, Center X, Director, Institute for Democracy, Education, and Access (IDEA), & Faculty, School of Education, UCLA
• Carmen Streigel, Director, ICT for Education and Training, RTI
• Jeremy Taylor, Director of Assessment and Continuous Improvement, CASEL
• Alexandra Teitel, School Director, PS/IS 284, New York City, New York
• Emma Vadehra, Consultant, former Chief of Staff to US Secretaries of Education Arne Duncan and John King
• Vic Vuchic, Chief Innovation Officer, Digital Promise Global
• Shirley Xu, English Teacher, Sharpstown High School, Houston, TX

Special Thanks To

• Mark Zelman
• Prashant Loyalka
• Sarah Berg
• Victor Orozco-Olvera
• Adam Zalisk
• David Kidd
• Olga Anisimova