

What Is the Potential for Applying Cost-Utility Analysis to Facilitate Evidence-Based Decision Making in Schools?

Fiona Hollands¹, Yilin Pan¹, and Maya Escueta¹

Education decision makers routinely make choices among programs and strategies to implement. Policy demands increasingly require that such decisions are based on evidence regarding program effectiveness at improving student outcomes. However, research evidence is but one of the considerations that practitioners must juggle, along with local conditions, capacity, resource availability, and stakeholder values. We investigated the feasibility of applying a multicriteria decision-making framework based on cost-utility analysis to facilitate evidence-based decisions by educators. Working with a total of 183 aspiring school leaders in class settings, we determined to what extent they could implement the initial steps of the framework. We subsequently invited three educators to apply the full framework to substantive decisions in their schools and report the results.

Keywords: decision making; descriptive analysis; program evaluation; qualitative research; research utilization

For years, the research community, research funders, and the federal government have been exhorting education practitioners and decision makers at all levels of administration—school, district, and state—to use research evidence and data to guide decisions about how best to educate our youth. The Every Student Succeeds Act (ESSA; 2015) defines evidence-based activities, strategies, and interventions and applies these standards to the use of federal education funds. Recognizing that a majority of educational interventions have not been studied rigorously, one variant of the ESSA definition allows, under some circumstances, for funds to be applied to interventions that have not as yet been rigorously assessed. The caveat in such situations is that the intervention must be examined on an ongoing basis to assess improvements in student or other relevant outcomes. This sets an imperative to provide education practitioners with the tools to engage in such assessments.

Beyond the fact that educational decisions must often be made when evidence is unavailable or inconclusive (Conaway & Goldhaber, 2018), a critical barrier to the use of research-based evidence is the tension between research evidence and ideology—values and preferences. Prevailing research methods that are intended to inform programmatic and policy decisions in education focus primarily on fidelity of implementation and impact on student outcomes. Occasionally, resource requirements are

considered in cost-effectiveness analyses, or long-term economic benefits are assessed in cost-benefit analyses (Levin & Belfield, 2015). But these methods still fail to accommodate values and interests that cannot realistically be removed from policymaking and school board deliberations that must “balance technical and public concerns” (Asen, Gurke, Connors, Solomon, & Gumm, 2013, p. 54). In practice, an educational strategy that is incompatible with local values may not be accepted even if it is documented to be effective. Research evidence may also not be accepted by decision makers if its conclusions are not supported by what Feuer (2015) terms *experiential evidence*, which arises from professional practice and experience. Tseng and Nutley (2014) similarly observe that “research findings suggest a direction of travel, but specific actions are negotiated locally” (p. 170); additional considerations include “local data analyses, organizational history, and practice experience.” Tseng and Nutley argue that decision makers must be able to “weigh and integrate different types of evidence” (p. 170) and that the evidence must be “adjudicated alongside values, interests, and local circumstances” (p. 173). Related arguments have been made by Weiss (1995) and Edwards, Guttentag, and Snapper (1975), who contended

¹Teachers College, Columbia University, New York, NY



that evaluation research needs a “usable conceptual framework and methodology that links inferences about states of the world, the values of decision makers and decisions” (p. 140).

The idea of considering multiple forms of evidence when making programmatic decisions in education, in addition to local values and judgments, is appealing in principle, primarily because it more nearly accords with how decisions are made in reality (Farley-Ripple, May, Karpyn, Tilley, & McDonough, 2018; Honig & Coburn, 2008). However, it is practically challenging to account for multiple dimensions of program value, often judged differently by each set of stakeholders, and subsequently to distill all dimensions and judgments into a single decision. Edwards et al. (1975), building on earlier work by Edwards (1971), propose a framework to operationalize such complex decision making based on the idea of “utility” or usefulness of social programs. Utility has been variously defined by different academic disciplines, but the notion essentially revolves around an individual’s satisfaction derived from a good or service. For example, utilitarian philosopher Jeremy Bentham (1789/1907) defined utility as the “property in any object, whereby it tends to produce benefit, advantage, pleasure, good, or happiness” (p. 1.4). Economists (e.g., von Neumann & Morgenstern, 1944) have developed sophisticated ways to calculate expected utility using a combination of decision-maker preferences and probabilities of outcomes, but these are difficult to apply in practice.

The decision-making framework proposed by Edwards et al. (1975) derives utility using multi-attribute utility theory (MAUT), one of a family of multicriteria decision-making methods that are widely applied in fields such as operations research, finance, business, management science, computer science, environmental science, economics, energy, engineering, public sector policymaking, and medicine (Belton & Stewart, 2002; Dolan, 2010; Kiker, Bridges, Varghese, Seager, & Linkov, 2005; Zopounidis & Doumpos, 2017). These methods aim to address the tradeoffs between multiple objectives in a complex decision context (Ishizaka & Nemery, 2013; Pan, 2016). MAUT provides a practical, step-by-step mechanism to transform values of attributes, or what we term *evaluation criteria*, measured using different scales and units into utility values, a uniform measure of preference. Subsequently, the utilities of different outcomes are combined quantitatively to assign an overall utility value to each alternative program choice (Abdellaoui & Gonzales, 2009; San Cristóbal Mateo, 2012).

Saar (1980) recommended MAUT as an evaluation model appropriate for use by practitioners in schools because it blends educators’ judgments with objective evidence, links evaluation to planning, can be conducted in a more timely and less costly manner than formal external evaluations, and engages local stakeholders in the evaluation. But despite the apparent fit between multicriteria decision analysis and the needs of decision makers in education, it has been applied in only a few instances (e.g., Blanchard, Pierce, & Hood, 1989; Lewis & Kallsen, 1995; McCartt, 1986). Levin and others have extended the multicriteria approach and demonstrated how utility can

be juxtaposed against resource requirements in a cost-utility analysis to help decision makers choose among educational programs (Bitsoi Largie, 2003; Fletcher, Hawley, & Piele, 1990; Levin, 1980, 1983; Levin & McEwan, 2001; Lewis, 1989; Lewis, Johnson, Erickson, & Bruininks, 1994; Pruslow, 2001). These studies demonstrate that the cost-utility framework can facilitate the consideration of multiple types of evidence in the decision-making process, ranging from external research on program effectiveness and costs to internal data collected by the educational institutions and local knowledge drawn from teachers, administrators, parents, and community members.

Building on these pioneering applications, we set out to investigate the potential for cost-utility analysis to be used more routinely by practicing educators as a process to guide how decisions are made. Specifically, we sought to answer whether school-based decision makers (a) can apply steps of the cost-utility framework to facilitate a choice among alternative strategies to address a decision problem and (b) believe the framework could be useful for decision making in their regular school practice.

Methods

The decision-making framework based on cost-utility analysis consists of multiple steps adopted from Edwards et al.’s (1975) MAUT approach for establishing the utility of alternative strategies. Although presented linearly, the process may involve some degree of cycling back and forth between steps before progressing to a conclusion:

1. Identify the problem to be solved and the related decision to be made.
2. Identify solution options, that is, alternative educational programs or strategies that can address the issue.
3. Identify stakeholders or representatives who will actively participate in the decision-making process.
4. Create a list of evaluation criteria.¹ These are standards or dimensions by which the satisfaction derived from each solution option is to be judged, for example, its impact on reading comprehension or how much training is required for teachers to implement it. The criteria are assumed to be independent of one another such that their values are additive.
5. Each stakeholder (or representative) assigns importance scores between 0 and 100 to each evaluation criterion to indicate the relative importance of each criterion. Scores are averaged across stakeholders and normalized to produce importance weights for each criterion that sum to 1.
6. Identify measures to assess the extent to which each solution option satisfies each evaluation criterion. Establish plausible low and high scores for each measure to determine minimum (0) and maximum (10) utility and graph utility functions. We adopt Edwards et al.’s (1975) assumption of linear utility functions.

7. Collect numerical data or judgements to evaluate each solution option against each evaluation criterion and use these to calculate criterion-level utility values.
8. Calculate overall utility of each solution option by summing the products of importance weights and criterion-level utility values.
9. Estimate the costs per participant of implementing each solution option using the ingredients method (Levin & McEwan, 2001).
10. Calculate cost-utility ratios and rank solution options; test robustness of ranking with sensitivity analyses and use the results to inform the decision.

We conducted our investigation in three parts. First, we introduced the framework to three cohorts of principals, assistant principals, teacher leaders, and teachers attending a principal preparation program at a university in the Northeast. During a 90-minute session for each cohort, we engaged them in Steps 1 through 6 of the framework to assess whether they could apply them to decisions about new schools they were designing. Second, we surveyed the educators in Cohort 2 at the end of the in-class activity to elicit their views on the value of the framework and whether they would want to apply it to their regular school practice. Finally, during 2017–2018, we invited three educators from Cohort 3 to collaborate with us on applying the full framework to a real-life decision at the schools in which they work.

The educators in the principal preparation program were a demographically diverse group from public and charter schools across the United States. They were working in small groups of four to six people to design new schools over several months. Cohort 1 met in July 2016 and consisted of 83 educators working in 15 groups. Cohort 2 met in January 2017 and consisted of 51 educators working in 13 groups. Cohort 3 met in January 2018 and consisted of 49 educators working in 9 groups. For each cohort, we collected and analyzed data on their ability to implement each of the first six steps of the decision-making framework. We first asked members of each group to identify and discuss among themselves a decision problem relevant to their new school design. Next, each group was asked to agree on two or three solution options to consider. Each group reported their decision problem and solution options to the class for initial feedback and recorded the decision problem, solution options, and all subsequent information electronically in a Qualtrics-based survey tool that we created to facilitate the implementation of the decision-making framework. In a third step, each group discussed who would constitute stakeholders that should be involved in the decision-making process and entered this list in the Qualtrics-based survey tool.

In a fourth step, each group agreed a list of evaluation criteria that group members would use to compare and judge the solution options and entered these into the survey tool. The tool promptly presented these same criteria to the users with an interactive drag feature allowing the users to assign an importance score out of 100 to each criterion. For the first two cohorts, following Edwards and Newman (1982), we instructed the educators to first identify the most important criterion and assign to it a score of 100. They were then asked to assign scores out of 100

to indicate how important each other criterion in their list was relative to the most important one. If they were all equally important, they could each be scored 100. This strategy aims to help decision makers recognize their priorities and score different criteria accordingly. However, many groups spent so long arguing about which criterion was most important that for Cohort 3, we eliminated this substep, allowing the group members to assign scores in any order. This speeded up the process as it was easier for group members to compromise about individual scores than about which criterion to single out as most important. In a sixth and final step, the groups were asked to describe a means to assess each solution option against each criterion, for example, through the collection of data or identification of existing research.

Using the data that each group entered into the Qualtrics-based tool, we determined how many groups successfully completed each step in the framework. We reviewed whether each group had identified methods to evaluate each solution option against their stated criteria and whether these methods were appropriate for the purpose. To complete a full cost-utility analysis, additional steps (7–10) would include collecting data to evaluate the solution options, transforming the results into utility values, estimating the costs of each solution option, and calculating cost-utility ratios. Time constraints during these 90-minute sessions precluded the collection of evaluation data that ideally involve multiple stakeholders and substantive data-gathering activities in a real-world setting.

For the second part of our data collection, we solicited survey feedback using Likert scale and open-ended questions from the educators in Cohort 2. Forty-one of the 51 educators completed the survey. The Likert-scale questions inquired whether the educators had learned more about evidence-based decision making during the session, whether what they had learned would influence their work in schools, whether they would like to incorporate evidence-based decision making in their leadership practice, and whether they would like to incorporate the online decision-making tool in their regular practice. Response choices were: *strongly agree*, *agree*, *neutral*, *disagree*, and *strongly disagree*. An open-ended question asked for the educators' thoughts on the value of cost-utility analysis as a tool for decision making in their schools.

In the third stage of our investigation, we invited 3 of the 49 educators from Cohort 3 to apply the full framework to a real-life decision at the schools in which they worked during 2017–2018. These individuals were recommended to us by the administrators of the principal preparation program.

Data and Results

Classroom Application of Framework

Across the three cohorts, the 37 groups varied in the amount of time taken to proceed through the first few steps of the cost-utility decision-making framework, with some finishing the steps easily within the 90-minute session and others still arguing at the end of the session about importance scores for each criterion or struggling to identify methods to evaluate each option against each criterion. These differences appeared to be related to

Table 1
The Number of Groups Successfully Completing Each of Steps 1 Through 6 in the Decision-Making Framework (Out of 37 Groups)

Step in Framework	N of Groups Completing This Step Successfully
1. Identify the problem to be solved and related decision.	37 ^a
2. Identify solution options.	37 ^b
3. Identify stakeholders.	35
4. Create a list of evaluation criteria.	35
5. Assign importance scores.	31
6. Identify measures to assess the extent to which each solution option satisfies each evaluation criterion.	18 ^c

^aIncluding one group that reported this step verbally but did not submit the information electronically.

^bIncluding two groups that reported this step verbally but did not submit the information electronically.

^cSix additional groups identified measures for some but not all of their evaluation criteria.

Table 2
The Frequency and Percentage of Stakeholder Types Mentioned by 5 or More of 35 Groups

Stakeholder	Frequency	% of Groups in Which the Stakeholder Was Mentioned
Teachers	30	86
Students	29	83
Parents/families	25	71
Community/district partners	12	34
Principal	11	31
Administrators	8	23
Instructional coach	7	20
Leadership team	6	17
Board members	5	14
Business office personnel	5	14

factors such as the complexity of the decision, how well aligned group members were in their visions for their new school, and group dynamics such as whether there was a clear leader who maintained forward momentum. Table 1 summarizes the number of groups that completed each step successfully.

Nineteen of the 37 decisions were related to instructional strategies, curricula, and tools, for example, deciding how to design a middle school, project-based curriculum, or deciding which dual-language math curriculum to adopt. The remaining 17 decisions were about logistical issues or school organization. Examples include where to locate a new school, which student information system to acquire, and how teachers would submit lesson plans. Of the 36 groups that electronically submitted information on their decision-making process, 35 provided either two or three solution options. For example, Group 1D was deciding between two dual-language math curricula and entered TERC Investigations and Eureka Math as the possible solution options. Each group listed between 3 and 10 stakeholder groups, with an average of 6.5. Table 2 lists the stakeholder categories mentioned by five or more groups and the percentage of these groups listing the stakeholder category. Teachers, parents/families, and students were the mostly commonly named stakeholders. For example, Group 1D, which was deciding on a dual-language math curriculum, listed the following as stakeholders: principal, dual-language coordinator,

students, parents, teachers, community coordinator, school board, and leadership team.

Each group listed between 2 and 11 evaluation criteria, with an average of 7.4. The most commonly listed criteria, mentioned by 19 groups each, were feasibility of implementation and impact on student performance. Table 3 shows the criteria identified by five or more groups. Thirty-one of the 35 groups that had listed evaluation criteria assigned importance scores ranging between 5 and 100 to each of the criteria to indicate the importance of each one relative to the others. It is not clear why the four remaining groups that had successfully provided a decision problem, stakeholders, and evaluation criteria failed to assign importance scores. Given the extensive discussions among group members on this issue, it is possible that they could not reach agreement among themselves in the time available.

Twenty-four groups of the 35 that had identified evaluation criteria were also able to articulate methods to evaluate each solution option against some or all of their criteria. Among these, 18 groups were able to identify evaluation methods for all criteria, whereas the remaining 6 groups provided evaluation methods for only some criteria. Again, it is possible that some groups did not have enough time to complete this task, but based on the questions raised during the class sessions and indications from some of the groups that they had done all they could, it appears that this step was particularly challenging. Even

Table 3
The Frequency and Percentage of Criteria Mentioned by 5 or More of 35 Groups

Evaluation Criterion	Frequency	% of Groups Mentioning Criterion
Feasibility of implementation	19	54
Impact on student performance	19	54
Cost of the intervention ^a	15	43
Teacher preference/buy-in	13	37
Alignment to school curriculum/standards	12	34
Alignment to school mission and vision	11	31
Culturally and contextually relevant/appropriate	11	31
Training/professional development needs	11	31
Teacher/staff workload	8	23
Level of student engagement	7	20
Student buy-in	7	20
Impact on student socio-emotional development	6	17
Teacher capacity	6	17
Parent buy-in/support	5	14

^aIn a full application of a cost-utility analysis, although being below a specific cost threshold might be included as an evaluation criterion, actual cost would be the numerator in the cost-utility ratio.

Table 4
Evaluation Criteria, Importance Scores, and Measures Used by One Group for Deciding on a Dual-Language Math Curriculum

Evaluation Criterion	Importance Score (0–100)	Measure
Student performance	100	Pre- and postassessments
Ease of implementation	70	Teacher survey, testimonials from other schools; curriculum analysis
Training required	70	Teacher survey; testimonials from other schools, curriculum analysis
Parent friendly	80	Parent surveys, testimonials from other schools
Standards alignment	90	Curriculum analysis
Home and school connection	80	Curriculum analysis
Quality of language/translations	90	Curriculum analysis
Differentiated supports	30	Curriculum analysis
Outside resources required	30	Curriculum analysis, testimonials from other schools

for the groups who identified methods of evaluation, participants tended to use the terms *research*, *rubric*, *data*, or *past projects* and other vague or general descriptions without specifying how to design rigorous methods for collecting data to inform their decisions. Table 4 provides an example of data collected from one group, presenting the evaluation criteria listed by Group 1D, along with importance scores that group members attached to each criterion and the evaluation method they articulated to assess each solution option against each criterion.

Survey Feedback on Cost-Utility Framework

Forty-one of the 51 educators in Cohort 2 responded to an end-of-session feedback survey. Eighty-eight percent of the respondents either strongly agreed or agreed with the statement “I would like to incorporate evidence-based decision making in my leadership practice.” Eighty-three percent of the respondents indicated that they had learned more about evidence-based decision making during the session. Seventy-one percent of the

respondents either strongly agreed or agreed that what they learned about evidence-based decision making using the cost-utility framework would influence their work at schools. Sixty-one percent reported that they would like to incorporate the online decision-making tool that we provided to facilitate decision making in their regular practice.

Thirty-five of the survey respondents provided an answer to the open-ended question, “What are your thoughts on the value of cost-utility analysis as a tool for decision making in your school?” Twenty-six of the respondents indicated that the framework would be important, useful, or effective in their school decision-making processes. Some were specific about the ways in which it would be helpful, for example, for making decisions based on “comprehensive information,” for getting “the most out of every dollar” when resources are limited, for managing the school budget, in making “optimal and research-based improvements,” and providing “a quantitative analysis on qualitative decisions.” Two people mentioned that the framework is helpful for considering stakeholders. Another two indicated that it

would be useful for convincing stakeholders about a decision or making information about decisions visually accessible to stakeholders, for example illustrating tradeoffs such as costs versus student achievement. Three respondents indicated that the framework helps in articulating what factors matter to them. Three respondents suggested that the framework would be most applicable for longer term or complicated decisions as opposed to day-to-day decision making. Five educators observed that applying the framework is time-consuming. One of these educators indicated that, as a result, he or she would use it sparingly, and another indicated that he or she would only use some parts of the framework. Three of the 41 respondents were not sure about how they would use the framework.

Ongoing Application of Cost-Utility Decision-Making Framework in Three Schools

In September 2017, we began our collaboration with the three individual educators who agreed to apply the cost-utility framework to a real decision at their own schools. We focus here on reporting the process of applying the framework as a demonstration of whether and how the framework is applicable and feasible in practice. We will report actual results of the cost-utility analyses separately in the future.

One educator, an assistant principal (AP) at a charter high school, focused on selecting the best option for providing instruction to students who are temporarily out of school for discipline or health reasons. The AP already had many ideas about how to address the problem, which she articulated as seven solution options. However, before applying evaluation criteria to select among them, she introduced an additional step to the framework that we termed a *screening criterion*. This was essentially a non-negotiable requirement that had to be met before a solution option could be further considered. In this case, the screening criterion was that the solution option must be legally compliant with state code. A careful review of the regulations governing instruction for students who are out of school for health or discipline reasons led the AP to conclude that two of the seven options should be eliminated without further consideration.

Two additional solution options were combined when it became apparent that neither one could serve both students out of school for health reasons and those out of school for discipline reasons. Four solution options remained, and the AP invited several instructional personnel to consider the viability, pros, and cons of each option. These stakeholders each provided handwritten answers in a template and assigned importance scores to four evaluation criteria: fit with school culture; implications for teacher workload; quality of the instruction and curriculum to be used; and simplicity of the process for distributing, collecting, and grading student work. Feedback from the stakeholders led the AP to add a new, fifth solution option. The AP devised strategies to evaluate each solution option against the four criteria and scored each option using information gathered from the stakeholders, time-use estimates from an analysis of the resource requirements to implement each solution option, and her own judgments and experiences

from past implementation of out-of-school instruction. We also worked with the AP to devise cost estimates for each of the five solution options to complete the information required to obtain cost-utility ratios for each option. Finally, we varied some of the inputs in sensitivity analyses to assess how these affected the cost-utility rankings. The AP shared results of the analysis with senior decision makers at the educational management organization that oversees her school and met with them to discuss whether and how to pursue one of the solution options.

At the second school, a K–12 charter school, the AP is choosing a social and emotional learning (SEL) curriculum to implement in an effort to reduce high suspension rates and conflict among students. Initially, he selected an assessment instrument that would allow him to document and justify the need for an SEL program at the school. He also searched for research reports on SEL programs to identify possible solution options that had shown evidence of effectiveness. He primarily relied on one report by Jones et al. (2017), which reviewed 25 different SEL programs, describing their characteristics, components, and implementation and summarizing evidence of effectiveness. He applied four screening criteria to narrow the list of options. He ruled out any program that was not standalone, did not serve all K–12 grades, did not provide at least 20 weeks of curriculum, and for which emotional processes constituted less than 40% of the curriculum. The list of options was eventually narrowed to two SEL programs with an additional option being for the school not to adopt any SEL curriculum at this time (the “status quo” option).

The AP identified five evaluation criteria to help choose among the options: feasibility of integration into the current schedule, feasibility of planning requirements, availability of professional development opportunities and support, the extent to which the program involves parents and other members of the school community, and the degree to which students enjoy participating in the curriculum. The AP presented the options and evaluation criteria to teachers, administrators, and school staff during weekly professional development meetings. These stakeholders added one new evaluation criterion: grade-level differentiation/rigor of SEL curriculum. In groups, they discussed pros and cons of each solution option. Subsequently, each person documented his or her individual views on the viability of each option in an online survey and assigned importance scores to the evaluation criteria. Later, a subset of these stakeholders met to evaluate the solution options against five of the six criteria, assigning each option a numerical score for each criterion.

To evaluate the options against the sixth criterion that related to student enjoyment of the curricula, the AP conducted in-class demonstration lessons from each of the two SEL curricula and asked students to rate how much they enjoyed them on a scale of 0 to 10. The average student rating served as the relevant evaluation score for each of the solution options. We combined the averaged importance scores from the adult stakeholders with the full set of evaluation scores and transformed them into utility values, assuming a simple additive model and a straight-line utility function. Finally, we used the ingredients method (Levin & McEwan, 2001) to develop cost estimates for each solution option. The AP will use the results of the analysis to seek approval from the school board to move ahead with the preferred option.

The last of the three educators is in a teaching and administrative position at a public district school. Despite apparent enthusiasm for the framework by several personnel at the school and their suggestions for a handful of decisions to which it could be applied, the school did not settle on a decision problem within the timeframe available for this study. Regulatory constraints imposed by the school district limit some of the decision-making power at the school, and at least one of the decisions they initially intended to work on was “made for them” by the central district office.

Discussion and Conclusion

We believe that the work with educators and schools reported here constitutes the first formal application of cost-utility analysis to decision making at the school level in American K–12 schools. The data from the three cohorts of educators to whom we introduced the framework indicate that the initial steps can be successfully implemented by education decision makers in helping them articulate, plan, and organize a decision process. The survey data from the second cohort indicated that a majority of the educators would like to incorporate the framework into their regular school practice. Penuel, Farrell, Allen, Toyama, and Coburn (2018) note the general appeal to education leaders of frameworks and specific guidelines for action. The cost-utility framework meets this description by providing clear steps for gathering and applying multiple forms of objective and subjective evidence to decision making. Learning how to use decision-making frameworks is particularly relevant to educators who are preparing to become school leaders, and we believe the introductory session we developed and implemented could be extended over several sessions during principal training programs to solidify and build their capacity to make evidence-based decisions that involve multiple stakeholders.

It appears, however, that educators have variable capacity to identify and collect evidence to inform their decisions: Although they are able to identify the issues that are important to them, they do not always know how to systematically gather data to support one program choice over another to satisfy these priorities. A key challenge for future development and application of the cost-utility framework is to guide decision makers in finding or gathering suitable evidence to evaluate educational strategies against a variety of objective and subjective criteria. Providing easy access to repositories of research or evidence-based educational resources may be helpful as many of the school-based educators were not familiar with sources such as the What Works Clearinghouse, Education Resources Information Center (ERIC), Best Evidence Encyclopedia, or edreports.org. It was notable that for both the real-life, school-based applications of the cost-utility framework, the APs were not able to find peer-reviewed research on the effectiveness of the solution options at yielding improvements in their outcomes of interest. As a result, the evidence used to make decisions was based more on professional judgements, internal surveys or interviews, consideration of feasibility of implementation, and fit with the school's culture or mission.

When existing research is lacking and where feasible, it may be wise for decision makers to conduct pilot studies of alternative solution options at their schools to assess which option is the

most effective at improving target outcomes. This requires the capacity to design and implement such studies and execute them swiftly enough to inform budget decisions. Introducing educators to tools such as RCT-YES and Ed Tech RCE Coach or providing them with templates and guidelines for collecting their own data through surveys, focus groups, or interviews, could improve the quality of their data analysis and internally produced evidence. However, school personnel need the time and incentives to undertake these activities. Honig, Venkateswaran, and McNeil (2017) also highlight the importance of internal leadership in encouraging the use of research.

Several lessons have been learned from the application of the cost-utility decision-making framework to three school settings. In practice, the process extends over several months and is regularly interrupted by the daily demands of operating a school. This raises the question, alluded to earlier by a few of the survey respondents, of the feasibility of routinely applying the method to make decisions in education. Clearly, the full method cannot be applied to daily decision making, but the first few steps could be used to systematize decision-making practices that, as one of our collaborating APs observed, already occur more inconsistently for many nontrivial decisions: articulating problems and the decision to be made clearly, considering multiple solution options, engaging key stakeholders, and considering multiple aspects of each solution option before settling on one to implement. The full method can be used to enhance existing evaluative processes that already require a significant investment of time, for example, when schools or districts engage in a Request for Proposals process for adopting a new curriculum or assessment. Additionally, all schools and districts spend a considerable amount of time and effort each year deciding which programs, strategies, and personnel to include in their annual budget. Using the cost-utility framework to establish a systematic and transparent process for justifying each budget item could help focus decision makers' attention on questions of how to substantively assess the contributions of each one. This would bring a more rational and evidence-based approach to the allocation of annual budgets while still allowing political considerations to be included.

Additional steps are needed in the framework to initially structure the problem because clearly articulating the need to be addressed requires effort and evidence gathering to appropriately frame the issue and justify attention to it. This observation comports with the arguments of Belton and Stewart (2010) that structuring problems carefully is a critical antecedent for successfully applying multicriteria decision analysis. Identifying solution options to solve the problem also requires a significant amount of time and information gathering. The application of screening criteria is helpful to narrow the list of possible solution options that need to be evaluated. The decision-making process is more iterative than linear, with the list of solution options changing as new information is surfaced and new evaluation criteria are added by stakeholders. This is consistent with decision-making processes described in other fields (Corner, Buchanan, & Henig, 2001). Further work is needed to evaluate whether our assumptions are reasonable, for example, that evaluation criteria are independent of each other and utility increases (or decreases) in a straight line.

The autonomy of the school affects the extent to which such a framework is useful: If most major decisions are controlled by a central office, there may be few suitable applications for the framework. It is also apparent that if the school has a board that reviews its budget, as in the case of charter schools, the board may be more focused on keeping overall expenditures under control than considering whether investments in individual programs or strategies will yield worthwhile educational returns. To be successful in identifying solution options that yield the highest stakeholder satisfaction, the participants in the decision process must be willing to admit and include the subjective criteria that influence their preferences in addition to objective factors. It is also critical to include as stakeholders any individuals who will determine whether a recommended solution option will be financed and implemented.

Finally, it is clear that the last few steps of the framework will need to be automated. These involve calculating utility values by integrating importance scores with quantitative data collected from the evaluation of each option against each criterion and ranking the solution options by cost-utility ratios. We are currently applying these lessons learned to inform the development of a free, online tool to facilitate implementation of the cost-utility framework by education decision makers and their stakeholders and automate these final steps. A longer term question of interest will be to assess whether such decision processes that involve stakeholders and are informed by various forms of evidence actually lead to better educational outcomes than decisions made less systematically.

NOTES

Work reported in this article with the January 2018 cohort of educators (Cohort 3) and the ongoing work to implement the full decision-making framework with three schools has been generously funded by the William T. Grant Foundation. In addition to the time and effort of the school personnel with whom we are collaborating, several other members of our institution have contributed to the most recent work: Henry Levin, James Corter, Atsuko Muroga, and Aasiya Kazi.

¹Note that different terms are used by other scholars. Edwards and Newman (1982) suggest that an *attribute* is something that at least some stakeholders care about enough that failure to consider it might lead to a suboptimal decision. Lewis, Johnson, Erickson, and Bruininks (1994) use *criterion* to refer to overarching goals or dimensions of a program and *attribute* to refer to “an observable (and measurable) characteristic or property of an alternative that allows one to assess the extent to which the alternative satisfies a criterion” (p. 85). Keeney and Raiffa (1993) differentiate between objectives, attributes, and goals.

REFERENCES

- Abdellaoui, M., & Gonzales, C. (2009). Multiattribute utility theory. In D. Bouyssou, D. Dubois, R. C. Pirlot, & H. Prade (Eds.), *Decision-making process: Concepts and methods* (pp. 579–616). London; Hoboken, NJ: ISTE; John Wiley & Sons.
- Asen, R., Gurke, D., Connors, P., Solomon, R., & Gumm, E. (2013). Research evidence and school board deliberations: Lessons from three Wisconsin school districts. *Educational Policy*, 27(1), 33–63.
- Belton, V., & Stewart, T. (2002). *Multiple criteria decision analysis: An integrated approach*. New York, NY: Springer Science & Business Media.
- Belton, V., & Stewart, T. (2010). Problem structuring and multiple criteria decision analysis. In M. Ehrgott, J. Figueira, & S. Greco (Eds.), *Trends in multiple criteria decision analysis* (pp. 209–239). Boston, MA: Springer.
- Bentham, J. (1907). *The principles of morals and legislation* (Reprint). Oxford, UK: Clarendon Press. (Original work published 1789)
- Bitsoi Largie, E. (2003). *A cost-utility analysis of alternative bilingual education programs in the state of New Mexico* (Unpublished doctoral dissertation). Arizona State University, Tempe, AZ.
- Blanchard, W., Pierce, G. A., & Hood, S. M. (1989). Multicriteria analysis: Managing complexity in selecting a student-information system. *Research in Higher Education*, 30(1), 21–29.
- Conaway, C., & Goldhaber, D. (2018). *Appropriate standards of evidence for education policy decision-making* (CEDR Policy Brief No. 04032018-1-3). Seattle, WA: University of Washington.
- Corner, J., Buchanan, J., & Henig, M. (2001). Dynamic decision problem structuring. *Journal of Multi-Criteria Decision Analysis*, 10(3), 129–141.
- Dolan, J. G. (2010). Multi-criteria clinical decision support: A primer on the use of multiple-criteria decision-making methods to promote evidence-based, patient-centered healthcare. *The Patient: Patient-Centered Outcomes Research*, 3(4), 229–248.
- Edwards, W. (1971). Social utilities. *Engineering Economist, Summer Symposium Series*, 6, 119–129.
- Edwards, W., Guttentag, M., & Snapper, K. (1975). A decision-theoretic approach to evaluation research. In E. Streuning & M. Guttentag (Eds.), *Handbook of evaluation research* (Vol. 1, pp. 139–181). Beverly Hills, CA: Sage.
- Edwards, W., & Newman, J. R. (1982). *Multiattribute evaluation*. Beverly Hills, CA: Sage Publications.
- Every Student Succeeds Act of 2015, Pub. L. No. 114-95 § 114 Stat. 1177 (2015–2016).
- Farley-Ripple, E., May, H., Karpyn, A., Tilley, K., & McDonough, K. (2018). Rethinking connections between research and practice in education: A conceptual framework. *Educational Researcher*, 47(4), 235–245.
- Feuer, M. J. (2015). Evidence and advocacy. In M. J. Feuer, A. I. Berman, & R. C. Atkinson (Eds.), *Past as prologue: The National Academy of Education at 50. Members reflect.* (pp. 95–101). Washington, DC: National Academy of Education.
- Fletcher, J. D., Hawley, D. E., & Piele, P. K. (1990). Costs, effects, and utility of microcomputer assisted instruction in the classroom. *American Educational Research Journal*, 27(4), 783–806.
- Honig, M. I., & Coburn, C. (2008). Evidence-based decision making in school district central offices: Toward a policy and research agenda. *Educational Policy*, 22(4), 578–608.
- Honig, M. I., Venkateswaran, N., & McNeil, P. (2017). Research use as learning: The case of fundamental change in school district central offices. *American Educational Research Journal*, 54(5), 938–971.
- Ishizaka, A., & Nemery, P. (2013). *Multi-criteria decision analysis: Methods and software*. Chichester, UK: Wiley.
- Jones, S., Brush, K., Bailey, R., Brion-Meisels, G., McIntyre, J., Kahn, J., . . . Stickle, L. (2017). *Navigating SEL from the inside out: Looking inside & across 25 leading SEL programs: A practical resource for schools and OST providers: (Elementary school focus)*. Cambridge, MA: Harvard University, Graduate School of Education. Retrieved from <http://www.wallacefoundation.org/knowledge-center/Documents/Navigating-Social-and-Emotional-Learning-from-the-Inside-Out.pdf>
- Keeney, R. L., & Raiffa, H. (1993). *Decisions with multiple objectives: preferences and value tradeoffs*. Cambridge, UK: Cambridge University Press.
- Kiker, G. A., Bridges, T. S., Varghese, A., Seager, T. P., & Linkov, I. (2005). Application of multicriteria decision analysis in environmental decision making. *Integrated Environmental Assessment and Management*, 1(2), 95–108.

- Levin, H. M. (1980). Teacher certification and the economics of information. *Educational Evaluation and Policy Analysis*, 2(4), 5–18.
- Levin, H. M. (1983). *Cost-effectiveness: A primer*. Newbury Park, CA: Sage Publications.
- Levin, H. M., & Belfield, C. R. (2015). Guiding the development and use of cost-effectiveness analysis in education. *Journal of Research on Educational Effectiveness*, 8(3), 400–418.
- Levin, H. M., & McEwan, P. J. (2001). *Cost-effectiveness analysis: Methods and applications* (2nd ed.). Thousand Oaks, CA: Sage Publications.
- Lewis, D. R. (1989). Use of cost-utility decision models in business education. *Journal of Education in Business*, 64, 275–278.
- Lewis, D. R., Johnson, D. R., Erickson, R. N., & Bruininks, R. H. (1994). Multiattribute evaluation of program alternatives within special education. *Journal of Disability Policy Studies*, 5(1), 77–112.
- Lewis, D. R., & Kallsen, L. A. (1995). Multiattribute evaluations: An aid in reallocation decisions in higher education. *The Review of Higher Education*, 18(4), 437–465.
- McCartt, A. T. (1986). Multiattribute utility models and the tenure process. *New Directions for Institutional Research*, 1986(49), 43–61.
- Pan, Y. (2016). *Essays on applying Bayesian data analysis to improve evidence-based decision-making in education* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses Database. (UMI 10182138)
- Penuel, W. R., Farrell, C. C., Allen, A. R., Toyama, Y., & Coburn, C. E. (2018). What research district leaders find useful. *Educational Policy*, 32(4), 540–568.
- Pruslow, J. T. (2001). A comparative cost and utility analysis of three models of service delivery for special needs students. *Journal of Education Finance*, 27(1), 535–565.
- Saar, S. S. (1980). The role of decision-analysis and Bayesian approach in planning and evaluating education programs. *Educational Evaluation and Policy Analysis*, 2(3), 63–72.
- San Cristóbal Mateo, J. R. (2012). *Multi-criteria analysis in the renewable energy industry*. New York, NY: Springer.
- Tseng, V., & Nutley, S. M. (2014). Building the infrastructure to improve the use and usefulness of research in education. In K. S. Finnigan & A. J. Daly (Eds.), *Using research evidence in education: From the schoolhouse door to Capitol Hill* (pp. 163–175). Cham, Germany: Springer.
- von Neumann, J., & Morgenstern, O. (1944). *Theory of games and economic behavior*. Princeton, NJ: Princeton University Press.
- Weiss, C. H. (1995). The four I's of school reform: How interests, ideology, information, and institution affect teachers and principals. *Harvard Educational Review*, 65(4), 571–593.
- Zopounidis, C., & Doumpos, M. (Eds.). (2017). *Multiple criteria decision making: Applications in management and engineering*. New York, NY: Springer.

AUTHORS

FIONA HOLLANDS, PhD, is the associate director and senior researcher at the Center for Benefit-Cost Studies of Education, Teachers College, Columbia University, 525 W. 120th Street, Box 180, New York, NY 10027; fmh7@tc.columbia.edu. Her research focuses on optimizing the use of resources in education by assessing costs as well as effectiveness, benefits, or utility of educational programs, practices, and strategies.

YILIN PAN, PhD, is a researcher at the Center for Benefit-Cost Studies of Education and in the Education Leadership Data Analytics research team at Teachers College, Columbia University, 525 W. 120th Street, Box 180, New York, NY 10027; yp2266@tc.columbia.edu. Her research focuses on evidence-based decision making in resource allocation at the state, district, and school levels.

MAYA ESCUETA, MPP, is a researcher at the Center for Benefit-Cost Studies of Education, Teachers College, Columbia University, 525 W. 120th Street, Box 180, New York, NY 10027; mme17@tc.columbia.edu. Her research focuses on rigorous evaluations of educational programs and policies and integrating the findings into education decision making.

Manuscript received July 24, 2018

Revisions received January 29, 2019, and March 31, 2019

Accepted April 22, 2019