Published in final edited form as:

Evid Policy. 2019 January ; 15(1): 103–123. doi:10.1332/174426418X15172392413087.

Just Google it: Measuring schools' use of research evidence with internet search results

Zachary Neal, Jennifer Lawlor, Jennifer Watling Neal, Kristen Mills, and Kathryn McAlindon
Department of Psychology, Michigan State University

Abstract

Measuring the use of research evidence (URE) by schools has become a central of education researchers. However, it has proven challenging due to low response rates, social desirability bias, and costly or time consuming data collection methods. To overcome these challenges and meet the needs of research focused on URE, this paper introduces a non-reactive archival measure: Archival Search of Use of Research Evidence (ASURE). ASURE counts references to research or evidence on a school's or school district's website to capture the extent of its rhetorical use of research evidence. After illustrating the collection of ASURE in all public school districts in Michigan (N = 595), we use data on these districts to show that ASURE is reliable and valid, and thus offers a promising new strategy for measuring URE in schools. We conclude by considering future steps for exploring ASURE not simply as a measure of URE in schools, but instead as a measurement strategy for assessing URE in a broad range of organizational contexts.

Keywords

measurement; knowledge use; schools; internet

Driven by federal legislation (e.g. No Child Left Behind Act of 2002, Every Student Succeeds Act of 2015) and calls from education researchers (e.g. Mayer, 2001; Slavin, 2002), public school administrators in the United States are expected to adopt evidence-based school programming. A similar push toward evidence-based practice in education has occurred in the United Kingdom (e.g. Brown, 2015; Durbin & Nelson, 2014; Nelson & O'Beirne, 2014), Australia (e.g. Blackmore, 2002), and Canada (Campbell et al., 2017). However, studies continue to question whether administrators are using research evidence to inform their decisions (Greenberg et al., 2003; Hallfors & Godette, 2002; Ringwalt et al., 2002; Ringwalt et al., 2009), and some have been critical of the push for evidence-based practice (e.g. Baker & Welner, 2012; Anderson & Herr, 2011). This led Durbin and Nelson (2014) to observe that "teaching is not currently an evidence-informed profession" (p. 1) and more recently for Malouf and Taymans (2016) to observe that whether education has transformed into an evidence-based field remains an open question. They addressed the first part of this question by "assess[ing] the availability of technically sound research on

education interventions” (p. 454), while here we take up the second part: *Are educators using it?*

Measuring attitudes toward, capacity for, and actual use of research evidence has become a central focus of education researchers. In the United States, the William T. Grant Foundation has sponsored research on the use of research evidence in youth settings since 2009, while in 2014 the National Center for Research in Policy and Practice and the Center for Research Use in Education were formed to investigate whether educators are using research to inform their decision-making. Similar efforts toward measuring the use of research evidence in educational settings have also emerged in the United Kingdom (e.g. Williams & Coles, 2007; Durbin & Nelson, 2014; Nelson, Mehta, Sharples, & Davey, 2017), Australia (e.g. Blackmore, 2002), Canada (e.g. Lysenko, Abrami, Bernard, & Dagenais, 2015), Germany (e.g. Demski, Ackeren, & Clausen, 2016), and the Netherlands (e.g. Schildkamp, Poortman, Luyten, & Ebbeler, 2017). Research use can occur in many different ways, including for instrumental purposes such as selecting a program with demonstrated effectiveness, as a conceptual tool that informs how programs are examined but may not impact which program is selected, and as a symbolic gesture that confers legitimacy on the decision to select a particular program (Weiss 1979, Estabrooks, 1999). However, regardless of the specific way that research evidence is used in schools, measuring it has proven challenging for a number of reasons. Educators’ busy schedules make it difficult to collect data directly from them, and more than a decade of promoting evidence-based programming has raised the risk of educators providing socially desirable responses. Moreover, even when data can be collected from educators as individuals, this provides limited information about research use by schools as institutions.

The Archival Search for Use of Research Evidence (ASURE) is an archival measure that aims to overcome these challenges and meet the needs of research focused on the use of research evidence. ASURE directly measures what we call the *rhetorical* use of research evidence: the invoking of research or evidence in communication to outside audiences (e.g. parents). However, when a person or organization engages in rhetorical use of research evidence, it is often to communicate about an actual instrumental, conceptual or symbolic use of research evidence. Thus, ASURE may also indirectly measure these other types of use. It differs from other measures in three important ways. First, whereas most other measures aim to assess the use of research evidence by individuals, ASURE is focused on use at the organizational-level, for example by a school, school district, or other education entity. Second, rather than attempting to directly capture instances of actual research use, ASURE instead aims to identify instances of organizations communicating about use to outside audiences (i.e. engaging in rhetorical use). Third, as an archival measure, ASURE is non-reactive and collecting it does not require direct interaction with organizations or people (e.g. schools or educators), and thus is both less prone to social desirability biases and places less burden on respondents. Moreover, because ASURE is derived from data that are readily accessible using internet search engine results, it can be used at large and nested scales (e.g. all schools in a district, and all districts in a state) and in near real-time.

This paper is divided into five sections. In the first section, we briefly review the meanings of “use of research evidence,” then in the second section we discuss some of the challenges

to measuring the use of research evidence in schools. In the third section, we introduce the ASURE measure as a possible solution to some of these challenges. In the fourth section, we use data collected from all public school districts in Michigan ($N = 562$) to present evidence of ASURE's feasibility and reliability as a direct measure of rhetorical use, and evidence of its validity as an indirect measure of actual use. We conclude in the final section with a discussion of the limitations of ASURE, and directions for its further development.

What is “Use of Research Evidence”?

As early as 1980, scholars have noted the difficulty of defining the use of research evidence and the related constructs of “knowledge utilization” and “evidence informed decision making,” leading Larsen (1980) to describe it as “one of the most pressing needs in the field” of knowledge utilization (p. 428). Many efforts to define these concepts trace initially to the mid-1970s, with separate strands focusing on *types* and *degrees* of use.

Rich (1975) and Caplan et al. (1975) both articulated two types of research use – *instrumental* and *conceptual* – in the same year, however they defined them in subtly different ways. For Rich (1975), instrumental use occurs in the short-term and “refers to use...of information that can be documented: it appears directly in a memo,” while conceptual use occurs in the long-term and “refers to influencing a policy-maker’s thinking about an issue” (p. 241–242). In contrast, for Caplan et al. (1975), instrumental use is “specific, task oriented”, while conceptual use is “general, broadly oriented” (p. 10). In 1977, Knorr instead juxtaposed instrumental use with *symbolic* use, which “refers to post hoc legitimations of decisions already taken” (p. 179). This triumvirate of types of research use – instrumental, conceptual, and symbolic – were ultimately joined together by Pelz (1978), Weiss and Bucuvalas (1980), and Beyer and Trice (1982), but continued to evolve. For example, Weiss (1979) referred to what had been called instrumental uses as *enlightenment*, and subdivided symbolic uses into those that are *political* (i.e. to legitimize action) and *tactical* (i.e. to deflect criticism or delay action). More recently, Estabrooks (1999) adopted a new set of terminology for the same set of types, replacing instrumental with *direct*, conceptual with *indirect*, and symbolic with *persuasive*, finding that although they are distinct types of research use, they are each aspects of a single underlying construct of research use.

While much of the theorizing and empirical research on the use of research evidence has focused on defining its types, others have also aimed to define degrees of use. Larsen (1980) distinguished *complete* use in which research findings are used in their original form, from *adapted* use that modifies findings to fit the user’s context, and *partial* use that relies only on those findings that are relevant or fit existing beliefs. Knott and Wildavsky (1980) offered a more complete seven-stage framework for understanding the degrees of use. In their framework, the most minimal use of research evidence is *reception*, in which policy and decision makers have merely received research evidence. Following reception, use may proceed through the stages of *cognition* (i.e. the research is read), *reference* (i.e. it changes thinking), and *effort* (i.e. it is used to persuade others). The process concludes with the *adoption* of the actions justified by the research, their *implementation*, and their *impact*. Although this framework focuses on degrees of use, there is also overlap with frameworks

that focus on types of use. For example, the reference stage mirrors conceptual use, while the effort stage mirrors symbolic use.

Each of these definitions of the use of research evidence raise the question: what counts as “research” or “evidence”? In much of the work reviewed above, research evidence is ambiguously defined, but from context implicitly refers to a kind of knowledge generated via the scientific method (i.e. via systematic data collection and analysis to answer a pre-defined question). Importantly, this work is clear that research is defined by the process, not the person who generated it: a university professor can generate research evidence, but so too can a practitioner. Under this definition, data by itself is not research because while data is “systematically collected and organized to represent some aspect of schools” (Schildkamp et al., 2017, p. 242), it may not be subject to formal analysis or have been collected for the purposes of answering a pre-defined question. Accordingly, the constructs of “data use” and “data based decision making” that often arise in the field of education are different from the use of research evidence, but nonetheless share some similarities (e.g. Anderson et al., 2010; Schildkamp et al., 2017; Wayman & Stringfield, 2006). For example, there are multiple types of data use, and data may be used in decision-making to varying degrees. Thus, past efforts to measure whether or to what extent educators are using data, and whether or to what extent they are using research evidence, may inform each other.

Challenges to measuring the use of research evidence in schools

Despite recent attention on the importance of the use of research evidence by schools, few quantitative measures have been developed to measure it. Often schools’ use of research evidence is assessed qualitatively (e.g. Coburn, Toure, & Yamashita, 2009), which can be valuable for understanding the processes involved, but can be difficult to employ in large-scale studies and for other researchers to replicate. Although quantitative measures of the use of research evidence have been developed in related settings like health care, these require adaptation for use in schools (e.g. Squires et al., 2011). In this brief review, we do not attempt to present a comprehensive review of the measurement of the use of research evidence, but rather to summarize the measurement challenges identified by existing reviews (e.g. Squires et al., 2011; Gagnon et al., 2014), with a particular focus on identifying the extent to which those challenges extend to measurement in schools.

First, nearly all measures are individual-level, assessing the extent to which individuals use research or characterizing individuals’ attitudes about or ability to access research. The pool of individuals varies in scope, from teachers alone (Nelson, Mehta, Sharples, and Davey, 2017; Schildkamp, Poortman, Luyten, & Ebbeler, 2017), to teachers and principals (Williams & Coles, 2007), to teachers, administrators, and staff (Waymans et al., 2016). A similar individual-level focus appears in other contexts as well, with some of the most widely used measures focused on individual social workers and case managers (e.g. Aarons, 2004; Rubin & Parrish, 2011; Palinkas et al., 2016). Such individual-level measures can be aggregated within organizations (e.g. Demski, van Ackeren, & Clausen, 2016), however this yields an indirect organization-level measure and can be problematic when there is variation among the individuals within an organization (O’Brien, 1990). Thus, while existing

measures can provide insight into educators' use of research evidence, they are less suitable for measuring schools' use of research evidence.

Second, nearly all measures of the use of research evidence are collected via self-report using a survey or structured interview. Unless specific steps are taken to avoid priming respondents, for example by not asking explicitly about research or evidence (Neal, Neal, Kornbluh, Mills, and Lawlor, 2015; Neal, Neal, Mills, and Lawlor, in press; Nelson, Mehta, Sharples, and Davey, 2017), these types of measures can introduce the risk of a social desirability bias. For example, Anderson, Leithwood, and Strauss (2010) sought to assess the use of research evidence using a four-item scale that included asking school principals (on a 6-point scale): "I rely frequently on research evidence in my decision making." In the current policy climate that promotes evidence-based practice, it is difficult to imagine a principal selecting the "disagree" or "strongly disagree" response option. Indeed, among 237 principals, the mean response on this item was 4.86, well above the scale midpoint and corresponding to "agree." Similarly, Williams and Coles (2007) found that only 2.7% of principals in the United Kingdom reported having a negative attitude about research. Self-report measures that invite socially desirable responses can inflate estimates and lead to ceiling effects that hamper subsequent statistical analysis.

Third, the majority of measures with evidence of validity are multi-item scales. Although scales allow the researcher to assess reliability, measure multiple dimensions of the use of research evidence, and examine the measurement's structure, their collection also requires more time (and when participant incentives are offered, money). Some of the scales developed for educators have more than 40 items (Schildkamp et al., 2017 has 61, Wayman et al., 2016 has 53, and Lysenko et al. 2015 has 43). Scales frequently used in social work are similarly long: the Evidence-Based Practice Process Assessment Scale has 51 items (EBPPAS; Rubin & Parrish, 2011), and the Structured Interview for Evidence Use has 45 items (SIEU; Palinkas et al., 2016). The length of these scales may limit their use in large populations (e.g. state- or nation-wide samples) or in populations with limited time (e.g. educators).

Finally, because many measures of the use of research evidence are designed to be collected directly from respondents, they raise concerns about low response rates and patterns of missingness. Most studies measuring the use of research evidence in schools did not report their response rates (e.g. Anderson, Leithwood, & Strauss, 2010; Lysenki et al., 2015; Schildkamp et al., 2017), while those that did reported fairly low response rates (11.1% in Williams & Coles, 2007). Response rates were slightly higher among social workers (37.1% for Rubin & Parrish, 2011), mental health practitioners (21.9% for Nelson & Steele, 2007), and government officials (35% for Landry, Lamari, & Amara, 2003), but still well below levels necessary to avoid non-response bias. The exception appears to be when such measures are collected from a captive student population for the purposes of evaluating training in the use of research evidence (94.9% for Bradley & Herrin, 2004; 86% for Johnston et al., 2003).

Developing an archival approach to measuring the use of research evidence in schools

Having encountered some of these challenges during the conceptualization and data collection stages of a larger study about the use of research evidence in public education (J. Neal et al., In Press; Z. Neal et al., In Press; Neal et al., 2015a; Neal et al., 2015b), we sought to develop a new measurement strategy that might provide some solutions or supplement existing measures. Actual instrumental (i.e. concrete application), conceptual (i.e. informing the thinking process), and symbolic (i.e. as a source of legitimacy) uses of research evidence are difficult to observe or detect directly, often because they are subtle and occur “behind the scenes.” However, schools’ engagement in these types of uses can leave behind indirect traces, including instances of schools communicating to outside audiences (e.g. parents, politicians) about having engaged in the uses of research evidence. This led us to explore the possibility of an archival measure of use of research evidence that focuses not on capturing instances of actual use, but instead on capturing the traces of use in schools’ communications (i.e. rhetorical use).

There are many different archival sources where a school’s use of research evidence might be reflected, including school board meeting minutes, school improvement plans, and teachers’ classroom materials. However, among the most readily accessible and public-facing archival sources is a school’s own website, including the web pages and documents hosted there. Drawing on this archival data source, the Archival Search for Use of Research Evidence (ASURE) measures a school’s rhetorical use of research evidence by identifying and counting references to using research or evidence on its website. It is collected by performing the following search using Google

```
allintext:("research based" | "evidence based") site:XXX
```

where XXX is replaced with a school’s top-level domain. This search will return a list of all web pages within the school’s website that contain the phrase “research based” *and/or* the phrase “evidence based” (case insensitive, with or without a hyphen), along with an approximate count of the number of pages meeting these criteria. Accordingly, ASURE scores take the form of a count variable, and thus range from 0 to, in principle, positive infinity, but in practice likely rarely exceed 1000. Figure 1 shows an example for South Lyon Community Schools, with the ASURE score circled.

The selection of keyword phrases is critical; too many, too few, or inappropriate keywords can lead to errors. We experimented with a large number of keyword phrases in different combinations, each time reading the pages returned by the search to determine whether they included references to the use of research evidence. For example, we found that the phrases “use of research evidence” and “evidence informed decision making” are rarely used and resulted in undercounting. Conversely, the keyword “research” is commonly used on pages that do not reference the use of research evidence and resulted in overcounting (e.g. on a page describing 4th grade students’ research projects). In the context of Michigan public schools investigated in the next section, we found that the keyword phrases in the search

above offered a balance between including relevant pages and excluding irrelevant ones (see Table 1).

This approach offers a potential solution to many of the challenges associated with existing measures of the use of research evidence reviewed above. By focusing on a school's website, it measures use of research evidence at the organizational rather than individual level, and as an archival measure, it is nonreactive and thus reduces the risk of obtaining socially desirable responses and eliminated burdens on the respondents. Additionally, because this measure can be collected electronically using publicly accessible websites, the cost of data collection is minimal, and the risk of non-response is limited to schools that do not maintain websites. Moreover, it may offer some additional advantages. First, it can be collected in near real-time, allowing researchers to collect baseline data or to characterize a sample of schools very quickly (c.f. Naven et al., in press). Second, it can be collected at large scales, for example, about all schools in a state (as illustrated in the next section) or country. Third, it can be collected at multiple, nested scales by modifying the domain specified in the search query, allowing researchers to characterize the use of research evidence not only in individual school buildings, but also in whole school districts, or in county- or higher-level administrative units (e.g. intermediate school districts).

While ASURE has the potential to offer some advantages, it is important to be clear about what it is and is not designed to measure. It directly measures schools' rhetorical use of research evidence, which we define conceptually as the invoking of research or evidence in communication to outside audiences (e.g. parents), and which ASURE operationalizes as website references to research and evidence. In the next section, we present evidence of ASURE's feasibility and reliability as a direct measure of rhetorical use. In contrast, ASURE does not directly measure the extent or quality of schools' use of research evidence. However, to the extent that rhetorical use and actual use co-occur (i.e. when schools use research, they also communicate about having used it), ASURE may provide an indirect or proxy measure of the extent to which a school engages in the use of research evidence. In the next section, we present evidence of ASURE's validity as an indirect measure of schools' use of research evidence.

Illustration and Validation: The use of research evidence in Michigan Public Schools

We initially experimented with ASURE by collecting it for approximately half of all public school districts and intermediate school districts in Michigan in October 2016 ($N = 268$). We then collected ASURE for all districts in Michigan in March 2017 to more rigorously investigate its reliability and validity as a measure of the use of research evidence ($N = 595$); results reported in this section refer to this full-scale data collection. ASURE could not be collected for 33 districts because they either did not maintain a website, or their website was not fully indexed by Google. Thus, we were able to collect ASURE for 562 districts (94.4%), which took approximately two hours, highlighting its feasibility even in relatively large settings.

Because ASURE is a single-item measure, the only form of reliability that is relevant is test-retest reliability, which assesses a measurement's stability over time. To evaluate ASURE's test-retest reliability, we compare ASURE scores in October 2016 and March 2017. The high correlation ($r = 0.887$, $p < .001$; see table 2) indicates that ASURE scores were stable over this six-month period and offers evidence that it exhibits test-retest reliability. More concretely, schools' ASURE scores do not fluctuate dramatically over short periods of time, but instead offer a time-stable measure of their rhetorical use of research evidence.

Figure 1 illustrates the distribution of ASURE scores among Michigan public schools in March 2017. Approximately 27% of districts' websites contained no pages referring to research or evidence ($N = 153$), and an additional 16% of districts referred to them on only one page ($N = 89$). However, the majority of districts in Michigan refer to research or evidence on two or more of their web pages ($N = 320$, 57%). On average, districts refer to research or evidence on 9.62 web pages, however this masks substantial variation in the measure ($s.d. = 23.18$). For example, some local public school districts referenced research or evidence on more than 100 separate web pages. These ASURE scores directly capture the frequency of a school's rhetorical use of research evidence when engaging with the public through its website. Although web pages are authored by individuals, the positively skewed distribution suggests that ASURE is not affected by social desirability and the associated ceiling effects.

Qualitative Validation

Table 1 presents qualitative evidence of ASURE's content validity and provides a window into exactly what ASURE is measuring by reporting, for school districts with low (1), medium (5), and high (28) values of ASURE, the context within which these keywords appeared on their websites. The website for Madison District Public Schools, located in suburban Detroit, contained only one page using either keyword – in an annual education report letter sent to parents and community members (A1) – and thus receives an ASURE score of 1. This may communicate an instrumental use of research to the public because it suggests the district is focusing teacher professional development on learning strategies that are based on research, but it may also communicate a symbolic use of research to the public because it communicates directly to parents the district's commitment to high-quality pedagogical strategies.

The website for Huron School District, located on the Southern edge of Detroit Metropolitan Airport, contained these keywords on five separate pages, and thus receives an ASURE score of 5. A reference to “research based” in the district's vision statement on its home page communicates a symbolic use of research evidence, conveying immediately to all website visitors that the district values research-based instructional practices (B1). Likewise, references to “research based” in materials promoting a tax increase communicates a symbolic use of research evidence to the public, in this case of a specifically political nature (B4 & B5). Notably, although the phrase “research based” is used in exactly the same context in B4 and B5, these are nonetheless distinct items on the website and therefore are counted separately to arrive at an ASURE score. In some sense, this “double-counts” references to uses of research evidence, but it captures the fact that the district has placed

such references in multiple locations throughout its website. Like Madison District, Huron District also uses the phrase “research based” in its annual education report letter to parents (B2). However, in this context, it may be viewed as a communication of conceptual use because thinking about research-based practices guides the district’s progress toward school improvement. Finally, in a page providing parents links to special education resources (B3), the National Dissemination Center for Children with Disabilities is highlighted as a source for “research based” information, which explicitly communicates instrumental use of research.

Finally, the website for Mattawan Consolidated Schools, located in rural Southwestern Michigan, contained the relevant keywords on 24 distinct pages, but receives an ASURE score of 28. When the number of pages containing the relevant keywords exceeds 10, the number reported by Google on the first page of search results (see figure 1) is an approximation. However, we found that these approximations are typically quite close to the actual page count, and as we demonstrate below, exhibits strong test-retest reliability. These 24 pages communicate several different types of the use of research evidence to the public, including instrumental (e.g. C2, C3, C4), conceptual (e.g. C9), and symbolic (e.g. C17–24). However, while most of the pages identified by ASURE in this district assess rhetorical uses of research evidence (i.e., the communication of the use of research evidence to the public), some do not. For example, the phrase “research based” appears in the title of an item in the bibliography of a curriculum manual (C7), and in a newsletter to describe a new component of the SAT exam (C10), but in neither case is the keyword used to refer to the use of research evidence.

Quantitative Validation

Although ASURE directly measures schools’ rhetorical use of research evidence, it may also offer an indirect or proxy measure of schools’ actual use of research evidence. To explore this possibility, Table 2 presents quantitative evidence of ASURE’s validity as a measure of the use of research evidence using correlation coefficients. The statistical significance of these coefficients was evaluated using Monte Carlo permutation tests to avoid increases in Type-I error associated with large sample sizes.

Validity is a more complex concept than reliability, and is composed of multiple aspects (Messick, 1995; AERA, APA, & NCME, 1999). Here, we focus on what Messick (1995) calls the external aspect, which can be established through “convergent and divergent evidence...as well as evidence of criterion relevance and applied utility” (p. 745). First, convergent evidence is a “pattern indicating a correspondence between measures of the same construct” (p. 746), and aims to demonstrate that a new measure is consistent with an already established measure. To provide evidence of ASURE’s convergence, we compare ASURE scores to scores on the Evidence Based Practice Attitudes Scale (EBPAS; Aarons, 2004), which is a widely used measure of the use of research evidence. As part of a related study conducted in 2015, a random sample of 200 Michigan school district superintendents were invited to complete two adapted versions of the openness and divergence subscales of the EBPAS (61% response rate, N = 122; J. Neal et al., In Press; Z. Neal et al., In Press; Neal et al., 2015a; Neal et al., 2015b). In the first version, they responded to items focused

on themselves (e.g. “I like to use new programs to help students”), while in the second version they responded to items focused on their district (e.g. “My district likes to use new programs to help students”). We find that ASURE is positively correlated with (i.e. converges on) both the self-focused ($r = 0.129$, $p = 0.157$) and district-focused ($r = 0.155$, $p = .090$) versions of the EBPAS. Notably, the correlation with the district-focused version is slightly higher, which is expected because ASURE is intended to be an organizational-level, rather than individual-level measure of the use of research evidence. Although the correlation coefficients are in the expected direction, neither is statistically significant, the implications of which we discuss below.

Second, discriminant evidence is a “pattern indicating a distinctness from measures of other constructs” (p. 746). Given its measurement, there is one construct in particular for which distinctness is most in jeopardy, and thus for which discriminant evidence is most needed: the overall size of a district’s website. Because ASURE is measured as a count of web pages, there is the risk that it simply measures the size of a district’s website, but not the extent of a district’s use of research evidence. To provide evidence of ASURE’s ability to discriminate these two potentially related constructs, we compare ASURE scores to the total number of web pages on each district’s website. Despite its statistical significance, the weak correlation ($r = 0.109$, $p = .010$) indicates that ASURE is not merely measuring website size, and thus offers evidence of its discriminant validity. Importantly, because ASURE is necessarily less than or equal to the total number of web pages on a given site, some minimal correlation between these two constructs is required, however it is smaller than most other correlations presented in Table 2.

Third, evidence of applied utility involves demonstrating a relationship between the measure in question and “criterion measures pertinent to selection, placement...or other accountability purposes in applied settings” (p. 746). To provide evidence of ASURE’s applied utility, we focus on two broad categories of such measures that are expected to be associated with schools’ use of research evidence: organizational capacity and student achievement. Domitrovich et al. (2008) and others have argued that “the amount and type of resources available to deliver evidence-based services in schools are important organizational-level factors to consider” when understanding schools’ implementation of evidence-based practices (p. 14; Owens, 2004; Ringeisen, Henderson, & Hoagwood, 2003). Because locating, understanding, and implementing research evidence is time consuming, time is one important resource to provide capacity for the use of research evidence, which can be limited in districts serving large numbers of students. Accordingly, we find that ASURE is negatively correlated with district enrollment, controlling for revenue ($r = -0.143$, $p = .002$). Because the implementation of evidence-based programs can also be costly, money and especially capacity for greater per-pupil spending is another important resource. We find that ASURE is positively correlated with district expenditure, controlling for district enrollment ($r = 0.181$, $p < .001$). Finally, simply having the autonomy and expectation to engage in using research evidence is an important, but often overlooked factor in capacity. In Michigan, Intermediate School Districts (ISDs) are county-level agencies that exist primarily to provide support services, including consultation on and assistance with research. We find that ASURE is positively correlated with whether or not a given entity is an ISD ($r = 0.252$, $p < .001$).

Whereas capacity to engage in the use of research evidence is an important precursor to the actual use of research evidence, student achievement is an important outcome. Promotion of the use of research evidence in schools is most often advocated as a strategy for improving learning outcomes. For example, in the United States, the purpose of the What Works Clearinghouse is to facilitate the use of research evidence for the selection of the most effective programs, while the purpose of the IES-sponsored National Center for Research in Policy and Practice is to explore “how research can best be used to improve school performance and student outcomes.” The Education Endowment Foundation and What Works Scotland have similar aims in the United Kingdom. Reflecting the effect of the use of research evidence on outcomes, we examined the relationship between ASURE and students’ academic proficiency as indicated by scores on the Michigan Student Test of Educational Progress (M-STEP), a standardized test introduced in 2015 to measure students’ progress toward state educational standards and used by all public schools in Michigan. We find that ASURE is significantly positively correlated with the percent of a district’s students that achieved proficiency in the areas of math ($r = 0.134$, $p = .003$), science ($r = 0.278$, $p < .001$), social studies ($r = 0.140$, $p = .002$), but not literacy ($r = 0.076$, $p = .098$).

Discussion

ASURE provides a direct measure of schools’ or school districts’ rhetorical use of research evidence, which refers to their communication with outside audiences about engaging in the use of research evidence. It also provides an indirect measure of schools’ actual use of research evidence. A school’s ASURE score directly conveys the number of pages on its website where research or evidence is referenced, but a more conceptual interpretation of individual scores is challenging without a frame of reference. That is, when looking at a single school’s score, it is difficult to say whether it is high or low, good or bad. Instead, it is more useful to interpret schools’ ASURE scores with reference to other schools’ scores. First, there is an important difference between schools with ASURE scores of 0 and those with non-zero ASURE scores: it distinguishes schools for which notions of research and evidence are sufficiently important to be mentioned when communicating with outside audiences from those where notions of research and evidence are not particularly salient. Second, among schools with non-zero ASURE scores, it may be possible to draw comparisons. For example, schools with higher ASURE scores invests more effort in, and perhaps places more value on, the use of research evidence, than schools with lower ASURE scores. However, as an indirect measure of use of research evidence, precise numerical comparisons may not be appropriate. For example, the difference between schools with ASURE scores of 1 and 2 may be more important than the difference between schools with ASURE scores of 50 and 51, despite the numerical differences being the same. Similarly, the difference between schools with ASURE scores of 500 and 1000 may be negligible, despite the substantial numerical difference.

As the text excerpts in Table 1 illustrate, ASURE captures school districts’ communications to the public about instrumental, conceptual, and symbolic uses of research evidence while rarely capturing instances that are not related to the use of research evidence, providing preliminary qualitative evidence for its content validity. At the same time, these qualitative data offer an opportunity to consider some of ASURE’s limitations. First, as a single-item

measure, it cannot be used to independently measure and distinguish school districts' communication about different types of uses of research evidence. Second, if a webpage has recently been redesigned, or if it is written in a non-HTML language (e.g. Java, Flash), search engines may not have indexed the contents of their pages, leading to coverage issues. Finally, this measure could be susceptible to false positives and false negatives. False positives include cases where a district uses phrases like "evidence-based" on their website, but is referring to something other than the use of research evidence (e.g. C10). Conversely, false negatives include cases where a district communicates information about the use of research evidence, but does not use the phrases specified in the search query on their website. In principle, these types of errors could be corrected by using the same measurement procedure, but by narrowing or broadening the keyword phrases, respectively. While our qualitative examination of search results suggests that false positives are rare, the frequency of false negatives is unknown.

The results presented in Table 2 provide preliminary quantitative evidence for ASURE's reliability as a direct measure of schools' rhetorical use of research evidence, and of its validity as an indirect measure of schools' actual use of research evidence. Specifically, they show that ASURE is stable over time, consistent with an existing measure of the use of research evidence, distinct from website size, and associated in expected ways with measures of district capacity and student performance. Some of these findings, especially those that provide somewhat weaker evidence, offer greater insight into ASURE and warrant closer consideration.

It might be surprising that ASURE is not more highly correlated with the total size of a district's website, and thus that ASURE does not require normalization by website size (e.g. expressed as a proportion of all pages). A more detailed look at district websites revealed that total website size is driven primarily by the site's architecture. For example, most district websites provide a staff directory, but some directories list all staff on a single page (resulting in a small website), while others list each staff member on their own contact page (resulting in a very large website). This helps explain the low correlation between ASURE and website size, and also why it is unnecessary and inappropriate to normalize ASURE by website size.

The positive correlations observed between ASURE and EBPAS offer some convergent evidence, but this may be viewed as weak evidence because these correlations are not statistically significant. However, the lack of statistical significance may be driven by the fact that the EBPAS is a self-report measure, and thus subject to social desirability biases and ceiling effects observed in many existing measures of the use of research evidence. Indeed, in this sample we observe evidence of ceiling effects ($M = 3.78$, $SD = 0.39$ on a 5-point scale). This could suggest that the lack of a significant correlation may be more closely associated with threats to the validity of EBPAS, than of ASURE.

The correlations between ASURE and student achievement on standardized tests were positive and significant for most subjects, but smaller and non-significant for literacy. This is surprising given the volume of research on literacy-focused programs and practices. Indeed, the What Works Clearinghouse lists nearly twice as many literacy-focused programs than in

any other focus. The non-significant correlation observed for literacy achievement may reflect that schools are exposed to a large pool of literacy-based programs, making the selection and implementation of evidence-based ones more challenging. It also represents an opportunity to reflect on what ASURE does and does not attempt to measure. ASURE is designed as a measure of schools' rhetorical use of research evidence, but it does not directly measure schools' (fidelity of) implementation of the findings of that research. For example, a school may use research evidence to select a particular literacy program, but subsequently be unable to fully implement the program due to other time, cost, or political barriers. More generally, however, the modest correlations observed for achievement across subject areas can be explained by the fact that schools' use of research evidence is only one of many factors that may influence students' achievement.

There are a number of ways that ASURE might be used in future research. First, ASURE provides a direct measure of schools' rhetorical use of research evidence, which can be used to better understand how different schools communicate with their outside audiences, and the extent to which they highlight research and evidence in those communications. Second, it provides a new indirect measure of the extent of schools' actual use of research evidence, which can be used to examine both potential predictors (e.g., district demographics, resources) and outcomes (e.g., student achievement, parental satisfaction) faster and in larger setting than are possible with existing measures. Third, it can be used in a complementary fashion alongside other strategies for investigating the use of research evidence. For example, ASURE might be paired with additional quantitative measures or qualitative investigation to identify specific types of use of research evidence, barriers to use, or the process of use. Such a combination of multiple measures and modes of investigation of the use of research evidence can provide a more comprehensive understanding of how schools are engaging with research. Finally, ASURE can be used as a diagnostic measure to identify schools that are or are not using research and evidence, or an evaluative tool to examine the impact of interventions intended to facilitate schools' use of research evidence.

However, supporting these future research applications, additional research on ASURE as a measure of the use of research evidence is needed. The evidence presented in Tables 1 and 2 of its reliability and validity should be viewed as tentative, and future work should aim to establish its reliability and validity in other contexts (e.g. other states) and at other organizational levels (e.g. school buildings). Additionally, further investigation of methods for selecting keyword phrases used in the search query may help identify ways to tailor these keywords. For example, while the phrases "research based" and "evidence based" are commonly used in primary public education settings in the United States, other phrases may be more appropriate for investigating the use of research evidence in other settings (e.g. higher education, public health) or locations (i.e. outside the United States). For example, the phrase "evidence informed decision making" (EIDM) is not widely used in public education settings in the United States, while it is more common in the field of public health in Canada (National Collaborating Centre for Methods and Tools, 2016).

More broadly, future research may also explore whether ASURE can be adapted as a general measure of organizations' rhetorical and actual use of research evidence. While in this paper we have described and validated ASURE in the context of public schools, in principle this

measurement strategy could be employed to measure the use of research evidence by any kind of organization that maintains a website. Again, because different kinds of organizations may use different keywords and phrases to refer to research evidence, the specific terms used in ASURE may vary across contexts. However, this initial exploration offers a starting point for the development not simply of a measure for schools, but instead of an archivally-based, organization-level measure of the use of research evidence.

Our brief review of existing strategies for measuring the use of research evidence in schools and related social service settings revealed a series of measurement challenges: a focus on individuals but not organizations, risks of social desirability bias and ceiling effects, time consuming and costly data collection, and low response rates. To address these challenges, we developed the ASURE measure, which uses readily available internet search engine results to directly measure schools' rhetorical use of research evidence, and indirectly measure their actual use of research evidence. Despite its potential limitations, our application of ASURE to measure the use of research evidence in all 595 Michigan public school districts suggests that ASURE holds promise as a rapid measure of use of research evidence that is both reliable and valid, and thus offers education researchers a viable new measurement strategy.

Acknowledgments

This work was supported by the William T Grant Foundation (182241, 183010) and the National Institute of Mental Health (R21 MH100238-01A1). It was approved as exempt by the Michigan State University Institutional Review Board (x12-1011e, x14-706e, x14-1173e).

References

- Aarons GA. 2004; Mental health provider attitudes toward adoption of evidence-based practice: The evidence-based practice attitudes scale (EBPAS). *Mental Health Services Review*. 6:61–74.
- American Educational Research Association, American Psychological Association & National Council on Measurement in Education. Standards for educational and psychological testing. Washington, DC: American Educational Research Association; 1999.
- Anderson G, Herr K. 2011; Scaling up “evidence-based” practices for teachers is a profitable but discredited paradigm. *Educational Researcher*. 40:287–289.
- Anderson S, Leithwood K, Strauss T. 2010; Leading data use in schools: Organizational conditions and practices at the school and district levels. *Leadership and Policy in Schools*. 9:292–327.
- Baker B, Wellner KG. 2012; Evidence and rigor: Scrutinizing the rhetorical embrace of evidence-based decision making. *Educational Researcher*. 41:98–101.
- Beyer JM, Trice HM. 1982; The utilization process: A conceptual framework and synthesis of empirical findings. *Administrative Science Quarterly*. 27:591–622.
- Blackmore J. 2002; Is it only ‘What works’ that ‘Counts’ in New Knowledge Economies? Evidence-based Practice, Educational Research and Teacher Education in Australia. *Social Policy and Society*. 1:257–266.
- Bradley P, Herrin J. 2004; Development and validation of an instrument to measure knowledge of evidence-based practice and searching skills. *Medical Education Online*. 9:15.
- Brown, C. *Leading the Use of Research and Evidence in Schools*. London: Institute of Education Press; 2015.
- Campbell C, Pollock K, Briscoe P, Carr-Harris S, Tuters S. 2017; Developing a knowledge network for applied education research to mobilise evidence in and for educational practice. *Educational Research*. 59:209–227.

- Caplan, N, Morrison, A, Slambaugh, RJ. The Use of Social Science Knowledge in Policy Decisions at the National Level. Ann Arbor, Mich: Institute for Social Research; 1975.
- Coburn CE, Toure J, Yamashita M. 2009; Evidence, interpretation, and persuasion: Instructional decision making at the district central office. *Teachers College Record*. 111:1115–1161.
- Demski D, van Ackeren I, Clausen M. 2016; Zum zusammenhang von schulkultur und evidenzbasiertem handeln – befunde einer erhebung mit dem “Organizational Culture Assessment Instrument”. *Journal for Educational Research*. 8:39–58.
- Domitrovich CE, Bradshaw CP, Poduska JM, Hoagwood K, Buckley JA, Olin S, Romanelli LH, Leaf PJ, Greenberg MT, Ialongo NS. 2008; Maximizing the Implementation Quality of Evidence-Based Preventive Interventions in Schools: A Conceptual Framework. *Advances in School Mental Health Promotion*. 1:6–28. [PubMed: 27182282]
- Durbin, B, Nelson, J. Why effective use of evidence in the classroom needs system-wide change. National Foundation for Educational Research; 2014. Jan, Retrieved on 27 September 2017 from <https://www.nfer.ac.uk/publications/99942/>
- Estabrooks CA. 1999; The conceptual structure of research utilization. *Research in Nursing and Health*. 22:203–216. [PubMed: 10344701]
- Gagnon MP, Attieh R, Ghandour EK, Légaré F, Ouimet M, Estabrooks CA, Grimshaw J. 2014; A systematic review of instruments to assess organizational readiness for knowledge translation in health care. *PLoS ONE*. 9:e114338. [PubMed: 25474622]
- Greenberg MT, Weissberg RP, O'Brien MU, Zins JE, Fredericks L, Resnik H, Elias MJ. 2003; Enhancing school-based prevention and youth development through coordinated social, emotional, and academic learning. *American Psychologist*. 58:466–474. [PubMed: 12971193]
- Hallfors D, Godette D. 2002; Will the ‘Principles of Effectiveness’ improve prevention practice? Early findings from a diffusion study. *Health Education Research*. 17:461–470. [PubMed: 12197591]
- Johnston JM, Leung GM, Fielding R, Tin KYK, Ho LM. 2003; The development and validation of a knowledge, attitude and behaviour questionnaire to assess undergraduate evidence-based practice teaching and learning. *Medical Education*. 37:992–1000. [PubMed: 14629412]
- Knorr, KD. Policymakers’ use of social science knowledge: Symbolic or instrumental. In: Weiss, CH, editor. *Using Social Research in Public Policy Making*. Lexington: Lexington Books; 1977. 165–182.
- Knott J, Wildavsky A. 1980; If dissemination is the solution, what is the problem? *Knowledge: Creation, Diffusion, Utilization*. 1:537–578.
- Landry R, Lamari M, Amara N. 2003; The extent and determinants of the utilization of university research in government agencies. *Public Administration Review*. 63:192–205.
- Larsen JK. 1980; Knowledge utilization: What is it? *Knowledge: Creation, Diffusion, Utilization*. 1:421–442.
- Lysenko LV, Abrami PC, Bernard RM, Dagenais C. 2015; Research use in education: An online survey of school practitioners. *Brock Educational Journal*. 25:35–54.
- Malouf DB, Taymans JM. 2016; Anatomy of an evidence base. *Educational Researcher*. 45:454–459.
- Mayer RE. 2001; Resisting the assault on science: The case for evidence-based reasoning in educational research. *Educational Researcher*. 30:29–30.
- National Collaborating Centre for Methods and Tools. Evidence-Informed Decision Making Checklist. Hamilton, ON: McMaster University; 2016. (Updated 01 September, 2017) Retrieved from <http://www.nccmt.ca/knowledgerepositories/search/237>
- Naven L, Inglis G, Harris R, Fergie G, Teal G, Phipps R, Stewart S, Kelly L, Hilton S, Smith M, McCartney G, Walsh D, Tolan M, Egan J. Right Here Right Now (RHRN) pilot study: testing a method of near-real-time data collection on the social determinants of health. *Evidence and Policy*.
- Nelson, J; O’Beirne, C. Using evidence in the classroom: What works and why?. National Foundation for Educational Research. 2014. Jan, Retrieved on 4 December 2017 from <https://www.nfer.ac.uk/publications/IMPA01/>
- Messick S. 1995; Validity of psychological assessment: Validation of inferences from persons’ responses and performances as scientific inquiry into score meaning. *American Psychologist*. 50:741–749.

- Neal JW, Neal Z, Kornbluh M, Mills KJ, Lawlor JA. 2015a; Brokering the research-practice gap: A typology. *American Journal of Community Psychology*. 56:422–435. [PubMed: 26310694]
- Neal JW, Neal ZP, VanDyke E, Kornbluh M. 2015b; Expediting the analysis of qualitative data in evaluation: A procedure for the rapid identification of themes from audio recordings (RITA). *American Journal of Evaluation*. 36:118–132.
- Neal JW, Neal ZP, Lawlor JA, Mills KJ, McAlindon K. What makes research useful for public school educators? *Administration and Policy in Mental Health and Mental Health Services Research*.
- Neal Z, Neal JW, Mills K, Lawlor J. Making or buying evidence: Using transaction cost economics to understand decision making in public school districts. *Evidence and Policy*.
- Nelson, J, Mehta, P, Sharples, J, Davey, C. Measuring teachers' research engagement: Findings from a pilot study. Educational Endowment Foundation; 2017. Mar, Retrieved on 27 September 2017 from <https://educationendowmentfoundation.org.uk/our-work/resources-centre/research-use-survey/>
- O'Brien RM. 1990; Estimating the reliability of aggregate-level variables based on individual-level characteristics. *Sociological Methods and Research*. 18:473–504.
- Owens, RG. *Organizational Behavior in Education: Adaptive Leadership and School Reform*. Boston: Pearson Allyn and Bacon; 2004.
- Palinkas LA, Garcia AR, Aarons GA, Finno-Velasquez M, Holloway IW, Mackie TI, Leslie LK, Chamberlain P. 2016; Measuring use of research evidence: The structured interview for evidence use. *Research on Social Work Practice*. 26:550–564. [PubMed: 27616869]
- Peletz, DC. Some expanded perspectives on use of social science in public policy. In: Yinger, M, Cutler, SJ, editors. *Major Social Issues: A Multidisciplinary View*. New York: Free Press; 1978. 346–357.
- Rich RF. 1975; Selective utilization of social science related information by federal policy-makers. *Inquiry*. 12:239–245. [PubMed: 126957]
- Ringeisen H, Henderson K, Hoagwood K. 2003; Context matters: schools and the 'research to practice gap' in children's mental health. *School Psychology Review*. 32:153–68.
- Ringwalt CL, Ennett S, Vincus A, Thorne J, Rohrbach LA, Simons-Rudolph A. 2002; The prevalence of effective substance use prevention curricula in U.S. middle schools. *Prevention Science*. 3:257–265. [PubMed: 12458764]
- Ringwalt C, Vincus AA, Hanley S, Ennett ST, Bowling JM, Rohrbach LA. 2009; The prevalence of evidence-based drug use prevention curricula in U.S. middle schools 2005. *Prevention Science*. 10:33–40. [PubMed: 19002583]
- Rubin A, Parrish DE. 2011; Validation of the evidence-based practice process assessment scale. *Research on Social Work Practice*. 21:106–118.
- Schildkamp K, Poortman C, Luyten H, Ebbeler J. 2017; Factors promoting and hindering data-based decision making in schools. *School Effectiveness and School Improvement*. 28:242–258.
- Slavin RE. 2002; Evidence-based education policies: Transforming educational practice and research. *Educational Researcher*. 31:15–21.
- Squires JE, Estabrooks CA, O'Rourke HM, Gustavsson P, Newburn-Cook CV, Wallin L. 2011; A systematic review of the psychometric properties of self-report research utilization measures used in healthcare. *Implementation Science*. 6:83. [PubMed: 21794144]
- Wayman JC, Stringfield S. 2006; Data use for school improvement: School practices and research perspectives. *American Journal of Education*. 112:463–468.
- Wayman, JC, Wilkerson, SB, Cho, V, Mandinach, EB, Supovitz, JA. *Guide to using the Teacher Data Use Survey*. Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Appalachia; 2016. REL 2017–166 Retrieved from <http://ies.ed.gov/ncee/edlabs>
- Weiss, CH, Bucuvalas, MJ. *Social Science Research and Decision-Making*. New York: Columbia University Press; 1980.
- Weiss CH. 1979; The many meanings of research utilization. *Public Administration Review*. 39:426–431.
- Williams D, Coles L. 2007; Teachers' approaches to finding and using research evidence: an information literacy perspective. *Educational Research*. 49:185–206.

Figure 1.
Example Google Search for ASURE

Figure 2.
Distribution of ASURE

Table 1

Example uses of research on district websites

#	District/Web page (ASURE score)	Keyword in context
A <u>Madison District Public Schools (1)</u>		
A1	Annual education report to parents & community members [PDF]	"We are dedicating our professional development time to studying best practice and research based strategies to improve student learning."
B <u>Huron School District (5)</u>		
B1	Home page	"Vision statement: Every day, every hour, all students will be engaged with outstanding teachers using research based high-quality instruction"
B2	Annual education report to parents & community members [PDF]	"During our School Improvement progress we are continually focusing on data-driven, research-based best practices in classroom."
B3	Special Education helpful links	"The center [National Dissemination Center for Children with Disabilities] that provides information to the nation on... research-based information on effective practices for children with disabilities."
B4	Mailer promoting "Countywide Enhancement Millage" [PDF]	"Invest in great teachers that build relationships and deliver quality, research based instruction"
B5	FAQ about Wayne County Enhancement Millage	"Invest in great teachers that build relationships and deliver quality, research based instruction"
C <u>Mattawan Consolidated School (28)</u>		
C1	About Us – Teacher Evaluation	"At Mattawan Consolidated School, we subscribe to the Center for Educational Leadership's (CEL) approach to teacher evaluation systems to gain research-based methods and instruments to..."
C2	Autism Awareness	"The rapid increase in the number of students with ASD entering public schools in Michigan has created a critical need for more in-depth training for educators in the use of evidence based practices."
C3	Michigan K-12 Standards: Science	"The proposed Michigan standards are built on this [National Research Council] research-based framework."
C4	Reading Intervention Programs	"This [One Minute Reader] at home version of the Read Naturally program will give you simple, research-based strategies to help improve fluency in your developing reader."
C5	Peer to Peer Program	"Peer to Peer is a peer-mediated instruction and intervention (PMII) that has emerged as an evidence-based practice in special education."
C6	Informative Writing [PDF, handout]	A five paragraph essay structure can be " research based : examines and conveys complex ideas through reliable sources and based on quality research."
C7	Language! tm The Comprehensive Literacy Curriculum [PDF]	Reference to: "Fletcher, J., & Lyon, G.R. (1998). Reading: A research-based approach. In W. Evers (Ed.), What's gone wrong in America's classrooms? (pp. 49–90). Stanford, CA: Hoover Institution Press."
C8	Transmittal Form for Mentor Teachers [PDF]	"...comment about what activity occurred in each area listed as you worked with your mentee... research-based teaching strategies..."
C9	School ADvance Administrator Evaluation System	"Dr. XXX worked with principal investigators, Dr. XXX and Dr. XXX of XXX University, under a grant funded by the Wallace Foundation, to study principal practices with an emphasis on evidence based decision-making."
C10	Mattawan Consolidated School District Newsletter, Spring 2015 [PDF]	The redesigned SAT includes "Scale from 200–800 for evidence based reading/writing"

#	District/Web page (ASURE score)	Keyword in context
C11	Mattawan Consolidated School District Newsletter, Back to School 2015 [PDF]	"This workshop [strategies to help with a developing reader] will provide parents with a number of evidence-based strategies they can use at home to help their children improve their decoding of words, reading fluency, and remembering what they have read."
C12	Board of Education Meeting Minutes	Revised student seclusion and restraint policy: "encourage the use of proactive, effective, evidence and research based strategies and best practices to reduce the occurrence of challenging behaviors, eliminate the use of seclusion and restraint, and increase meaningful instructional time for all students..."
C13	Welcome to Second Grade: Information Packet for Families [PDF]	"Since I am a graduate student, I will continually be involved in research-based teaching this year."
C14	Second grade teacher webpage	"Since I am a graduate student, I will continually be involved in research-based teaching this year."
C15	Agreement between Mattawan Consolidated School Van Buren and Kalamazoo Counties AND Kalamazoo County Education Association, 2013–2016 [PDF]	"...the district agrees to...encourage the exploration, staff development, and utilization of a variety of successful teaching methods, including research based best practices."
C16	Agreement between Mattawan Consolidated School Van Buren and Kalamazoo Counties AND Kalamazoo County Education Association, 2016–2020 [PDF]	"...the district agrees to...encourage the exploration, staff development, and utilization of a variety of successful teaching methods, including research based best practices."
C17	Annual education report to parents & community members, 2016 [PDF]	"The core curriculum is implemented daily through the use of research based materials and instructional best practice/quality instruction."
C18	Annual education report to parents & community members, 2017 [PDF]	"The core curriculum is implemented daily through the use of research based materials and instructional best practice/quality instruction."
C19	Annual education report to parents & community members (draft), 2017 [PDF]	"The core curriculum is implemented daily through the use of research based materials and instructional best practice/quality instruction."
C20	Annual education report for Mattawan Later Elementary, 2013 [PDF]	"The core curriculum is implemented daily through the use of research based materials and instructional best practice/quality instruction."
C21	Annual education report for Mattawan Later Elementary, 2014 [PDF]	"The core curriculum is implemented daily through the use of research based materials and instructional best practice/quality instruction."
C22	Annual education report for Mattawan Later Elementary, 2017 [PDF]	"The core curriculum is implemented daily through the use of research based materials and instructional best practice/quality instruction."
C23	Annual education report for Early Elementary School, 2014 [PDF]	"The core curriculum is implemented daily through the use of research based materials and instructional best practice/quality instruction."
C24	Annual education report for Mattawan High School, 2017 [PDF]	"The core curriculum is implemented daily through the use of research based materials and instructional best practice/quality instruction."

Table 2

Correlations with ASURE

Measurement Characteristic	Variable (N)	Pearson correlation
Reliability	ASURE in October 2016 (268)	0.887 ^{**}
Convergent evidence	EBPAS for self (122)	0.129
	EBPAS for district (121)	0.155 ⁺
Discriminant evidence	Total number of pages in website (562)	0.109 [*]
Evidence of Applied utility: Capacity	Total enrollment, controlling for revenue (493)	−0.143 ^{**}
	Expenditure, controlling for total enrollment (493)	0.181 ^{**}
	Intermediate School District (562)	0.252 ^{**}
Evidence of Applied utility: Achievement	% proficient in literacy (472)	0.076 ⁺
	% proficient in math (482)	0.134 ^{**}
	% proficient in science (486)	0.278 ^{**}
	% proficient in social studies (486)	0.140 ^{**}

^{**}
p < 0.01,

^{*}
p < 0.05,

⁺
p < 0.1 (computed using 10000 Monte Carlo permutations)