WORKING PAPER 65

MATHEMATICA Policy Research

Working PAPER

BY STEVEN GLAZERMAN, IRA NICHOLS-BARRER, JON VALANT, JESSE CHANDLER, ALYSON BURNETT

Nudging Parents to Choose Better Schools: The Importance of School Choice Architecture

November 2018

ABSTRACT

We conducted a randomized factorial experiment to determine how displaying school information to parents in different ways might affect what schools they choose for their children. In a sample of 3,500 low-income parents of school-aged children, we found that a small nudge, such as changing the default order in which schools were presented, could induce meaningful changes in the types of schools selected. Specifically, changing the default sort order from distance-from-home to academic performance resulted in parents choosing schools with higher academic performance. The academic performance of the average school selected was 5 percentile points higher, equivalent to 0.20 standard deviations. The change in sort order also led parents to choose schools that were more than half a mile farther from home (2.3 versus 1.7 miles, on average). Other design choices such as using icons to represent data, instead of graphs or just numbers, or presenting concise summaries instead of detailed displays, also led parents to choose schools with higher academic performance. We also examined effects of information display strategies on parents' understanding of the information and their self-reported satisfaction and ease of use. In some cases, there were trade-offs. For example, representing data using only numbers maximized understanding, but adding graphs maximized satisfaction at the expense of understanding.

Many people contributed to this study. Eric Johnson provided important design advice early in the project. Phil Gleason read drafts and provided helpful comments. Daniel Kassler provided expert research assistance and Mariel Finucane and Shira Mitchell provided statistical consulting. The team at Tembo, Inc., led by Meg Towle and Michael Moore, designed and produced the school shopping site mockups that were used in the study. The study was funded by the U.S. Department of Education under Contract No. ED-IES-15-C-0048. The authors also received funding from the Walton Family Foundation. Neither funder is responsible for the content of the paper nor the views expressed herein.

I. INTRODUCTION

Every year the number of families that can choose what school their child attends increases, both because school districts are increasingly offering parents a choice between existing schools and because of the rise in alternatives to traditional public schools (National Center for Education Statistics, 2018; Whitehurst, 2017). For school choice to be good public policy, choosers—typically parents—must be well informed.[1] Moreover, individual choices can drive collective outcomes. If what parents desire from their children's schools aligns with what the public wants from a publicly funded education system, then school choices, taken together, could generate pressure for schools to serve societal interests such as academic excellence or racial and class integration. Otherwise, school choice could exacerbate inequality and drive a race to provide private amenities that have little social benefit.

One way that consumers in the education market become informed is through online displays of school choice information, sometimes referred to as school shopping websites or school finders. In some cities, like Washington, D.C., and New Orleans, the school finder is integrated with the official online portal for choosers to select schools and submit a common application, providing rank-ordered lists of desired schools.

The design of those sites can be important. Someone must determine the order in which to list the schools, the types of data to present, the layouts, fonts, level of detail, website navigation, tools for sorting and filtering, and other factors. Those who design the sites can be thought of as "choice architects," the term Thaler and Sunstein (2008) use for those who present information about choices. Choice architects can play a powerful role in shaping the decision process and, by extension, the outcomes of the system. Design can determine which choices or choice features are likely to be seen, against what standard or comparison a choice is evaluated, what strategy is used to evaluate choices, and how much effort families invest in choosing. All of these features determine which schools parents are likely to consider and whether they are likely to select any particular school.

A growing literature in education, reviewed below, is built on insights from the behavioral sciences and examines how policymakers can improve school shopping websites and, in turn, the functioning of school choice markets. Nevertheless, there is ample room for research to inform school choice architecture. A review of 14 school shopping websites across the United States revealed an assortment of design and presentation approaches with no consensus on how to present information to school choosers (Glazerman, 2017). Many basic, important questions remain unanswered about how best to design and present information to school choosers. This study presents evidence from an online experiment that aims to address this gap.

II. PREVIOUS RESEARCH

The current study builds on two strands of literature. One strand is about how parents choose schools. The other is a more general body of work to understand consumer choice and the ways

[1] For simplicity, this paper refers to *choosers* and *parents* interchangeably, although school choosers can include guardians, caretakers, other adult family members, and students themselves.

that information design influences consumer behavior. We discuss some of this research to motivate the current study.

A. Insights about how parents choose schools

Parents' values. Within the field of education, researchers have primarily focused on discovering what parents value in schools, rather than on how best to present school choice information. The most common approach has been to conduct focus groups or surveys that ask participants about the factors that drive their choices (Collins & Snell, 2000; GreatSchools, 2013; Kelly & Scafidi, 2013; Klute, 2012; Jochim, DeArmond, Gross, & Lake, 2014). Parents usually cite academic factors as the most important consideration when selecting a school (Bosetti, 2004; Teske & Schneider, 2001; Stein, Goldring, & Cravens, 2010), but studies that rely on revealed preference by observing choices suggest they value school safety, geographic location, and extracurricular activities (Collins & Snell, 2000), as well as the composition of the students in the school, including their proficiency rates, race, and ethnicity (Harris & Larsen, 2015; Glazerman & Dotter, 2017).

Low-income and low-information parents. Another line of research suggests that low-income choosers face a steeper challenge. Teske and Schneider (2001) found that parents with higher levels of educational attainment are more likely than their less educated peers to choose an alternative to the neighborhood school for their children. Schneider, Teske, Marshall, and Roch (1998) determined that on average, low-income parents had less objective knowledge about schools than higher-income parents.

Parents' social networks may also contribute to low-income families choosing different schools than higher-income peers (Neild, 2005). A study by Hastings, Kane, and Staiger (2009) examined parents' rankings of their top school choices and found that the preference for a school with high test scores increased with family income. Bell (2009) found that lower-income parents tend to have greater social ties with families using lower-performing schools, making them more likely to choose these schools. Bosetti's (2004) research likewise shows that social circles of more highly educated parents are more likely to include professionals who are knowledgeable about the system. These differences in social networks likely contribute to the information gap between lower- and higher-income parents.

School choice presentations. Compared with the research on parents' values and knowledge about schools, evidence on the presentation of school choice information is relatively rare. Experimental evidence demonstrates that providing information influences choice (Corcoran, Jennings, Cohodes, & Sattin-Bajaj, 2018; Hastings & Weinstein, 2008), but these studies did not systematically vary the way that information is presented. Jacobsen, Snyder, and Saultz (2014), who studied the effect of information formats on parents' perceptions of schools, found that using simplified letter grades had a greater influence on parents' school ratings than performance index ratings, proficiency percentages, and achievement labels; parents who viewed the letter grade format perceived greater differences between schools than those who viewed the other formats. More recent studies have estimated the impact that information presentation has on school choice attitudes and behavior. Valant (2014) used quick-turnaround online experiments and a regression discontinuity design to examine how parents update their opinions of local public schools after receiving various types of information. Brief positive or negative comments from local parents strongly affected respondents' opinions. Numerical government ratings also affected their opinions, but more modestly. Valant found that respondents were attracted to the source (parents) and style (narrative) of the comments, prioritizing parent survey ratings and comments to government ratings. Although these findings begin to address the issue of choice architecture, critical gaps remain in the literature on how best to present school choice information.

B. Research on consumer choice and information design

Researchers in cognitive psychology and related fields have explored the extent to which different types of information displays affect information processing and ultimately decision making. Much of the research focuses on the limited capacity of decision makers to consider all options and all available information about each option (Peters, Dieckmann, Dixon, Hibbard, & Mertz, 2007). Decision makers often look for shortcuts or heuristics to simplify the task of processing information and truncate the information search and decision processes when they find a choice that is good enough for their standards.

The present study extends this research to the area of school choice by testing how various ways of presenting information influence the understandability and usability of school information displays, and examining how these factors affect parents' ranking of schools. It is impossible to construct a school shopping site without making innumerable explicit and implicit design decisions. The potential consequences of all of these decisions are beyond the scope of a single study. We examine the impact of five design factors that vary across school shopping sites and were identified as of particular relevance through practitioner interviews and review of existing research. We describe the research and potential impact of these factors below.

Data format. School data can be expressed in various ways including simple numbers or percentages, charts, and letter grades or other icons that summarize and express judgments about school performance or group continuous values into categories. Normatively, numeric information is superior because it provides more precise information than either graphs or icons. In practice, however, visual aids might help people simplify, organize, and interpret information. People usually prefer more visual displays of information (Campos, Doxey, & Hammond, 2010), and under some circumstances, visual displays can improve decision making by structuring information so that it parallels the task at hand (DeSanctis, 1984; Vessey, 1991). In particular, graphs are generally considered superior to numeric information when making comparisons of relative value (Feldman-Stewart, Kocovski, McConnell, Brundage, & MacKillop, 2000; Vessey, 1991).

Icons, such as stars or letter grades, collapse alternatives into larger categories, simplifying choice, with the trade-off of losing information (Jacobsen et al., 2014). They increase the apparent difference between the best and worst alternatives by ensuring that they are clearly the upper and lower bounds of the range of possible alternatives (Parducci, 1965; Stewart, Chater, & Brown, 2006). Icons also enable people to evaluate otherwise difficult-to-understand metrics (Hibbard, Slovic, Peters, & Finucane, 2002). Icons have successfully communicated information to consumers in a variety of domains including food safety (Jin & Leslie, 2005); health care (Hibbard et al., 2002); nutrition (Borgmeier & Westenhoefer, 2009; Jones & Richardson, 2007); and education (Jacobsen et al., 2014).

Reference point: displaying district averages. The desirability of a product depends not just on its own attributes, but also the reference point or standard against which it is evaluated (Schwarz & Bless, 1992; Tversky & Kahneman, 1991). A consumer deciding from a list of alternatives is likely to compare each new choice with the previously considered alternatives. This is suggested by the research of Suk, Lee, and Lichtenstein (2012), who examined the role of sort order on choice, but this tendency also means that consumers often make choices based on incomplete information. Providing a calculated average for all schools in the district may eliminate these difficulties by offering a clear and consistent reference point that parents can use to evaluate all alternatives.

Source of information. When choosing schools, parents can draw upon information from both official sources (such as state test scores) and from other parents (such as word-of-mouth reports from other parents; Valant, 2014). Some information displays include parents' opinions of schools—most often as survey results—while many others do not. Consumers may prefer to make decisions based on assessments from peers because they are easier to map onto their own likely experience than are technical specifications or performance measures (Chen & Xie, 2005). For example, in education, schools usually measure academic proficiency through performance on standardized tests, but parents may feel this is an imperfect assessment of academic performance relative to a holistic impression of parents familiar with the school.

Amount of information. Creating school profiles requires difficult trade-offs in presenting enough information that parents learn about individual schools without presenting so much information that they become overwhelmed or confused. Previous research outside of education indicates that limiting the amount of available information can improve the understandability of a presentation (Cronqvist & Thaler, 2004), yet at the same time make choosers feel less satisfied with their choices (Bundorf & Szrek, 2010; Chakraborty et al., 1994). An alternative is to give parents some degree of control over how much information they see by providing a minimal amount of information that can be expanded with a click to reveal more detail (progressive disclosure; Loranger & Nielsen, 2006).

Default sort order. Another design factor is the order in which schools are presented. Even if users of an online display can re-sort the list, there must be a default. The tendency for people to stop considering choices once they have found one that is good enough for their standards leads earlier alternatives to be preferred to later alternatives (for an overview, see Bar-Hillel, 2015). Consumers also may find it easier to compare schools along the dimension by which they are sorted (Quaschning, Pandelaere, & Vermeir, 2014).

III. STUDY DESIGN, DATA, AND METHODS

A. Design of the experiment

The study was conducted as an online randomized experiment. We created a data set of school information representing a hypothetical school district designed for the study, and showed this information to parents of school-aged children using a variety of different website designs. We measured the impact of design decisions on three categories of outcomes:

1. Behavior: choosers' selection of schools

2. Knowledge: choosers' ability to process school information accurately
3. Attitudes: choosers' perceptions about whether the information is easy to use and satisfying

This was a factorial experiment, a study design that varies several factors at once to simultaneously test each one (that is, each type of manipulation strategy for the displays) using a single sample of study participants. The study tested five factors:

1. **Format** varied whether school information was shown in one of the following ways:
    a. Only using numbers
    b. Combining numbers with color-coded, A-F letter grade icons (for measures related to academics and safety)
    c. Combining numbers with horizontal bar graphs (for measures related to academics and safety)
2. **District average** varied whether the displays included the district's average value as a benchmark for key measures of school performance.
3. **Source of information** varied whether the displays included parents' opinions (numerical satisfaction ratings) as an additional source of information for measures related to academic quality and school safety.
4. **Amount of information** varied the total amount of information shown as one of the following:
    a. Low (4 measures per school)
    b. High (22 measures per school)
    c. Low with progressive disclosure, meaning that the amount of information was low by default with a user option to click and reveal all the additional information in the higher-information version of the display
5. **Default sort order** varied whether the schools were initially sorted by distance from home versus initially sorting by academic quality. In all displays, users could re-sort the list by distance, academics, or alphabetically, although the index number appearing next to each school (and shown as location markers on the district map) continued to reflect the default sort-order ranking of schools.

The experiment featured 72 different school information displays, representing every possible combination of the 12 levels of the study's five factors (3 x 2 x 2 x 3 x 2). Figure 1 illustrates how each of the factors was implemented in the school information websites shown to study participants, together with the features of the information display that were held constant.

Figure 1. Example of an information display tested in the experiment

What was common to all information displays:

Title banner with link to glossary and option to display or hide map

Map with home icon and pins representing locations of 16 schools

Banner with "Sort By" dropdown list and labels for each category of information (with question mark icons for more info)

What varied across information displays:

Factor 1: Format. Whether data were presented as numbers only, numbers + icons, or numbers + graphs.

Factor 2: District average. Whether display included row with district-wide averages.

Factor 3: Data sources. Whether display contained only district indicators or both district indicators and parent ratings.

Factor 4: Amount of information. Whether profiles display lower level of information, higher level of information, or lower info that can open to higher info.

Factor 5: Default sort order. Whether schools were initially sorted by distance to school or academic performance.

Note. This figure illustrates one of the 72 information displays prepared for the experiment. The right side shows a school display with the following levels of each factor: school data formatted as graphs, inclusion of a district average as a reference point, inclusion of parent ratings, progressive disclosure for the amount of information, and default sort by distance from home to the school. The boxes on the left side indicate how each factor modifies the display.

The research team used an online survey to conduct the experiment, with a final sample size of 3,500 study participants who reported an annual household income of less than $40,000 and at least one child ages 3 to 18. The sample was recruited from a national consumer panel maintained by GfK, a market research firm. The sample was not statistically representative of parents nationwide because respondents opted in to the survey, but respondents were located in all 50 states and represented urban, suburban, and rural areas. Table 1 shows the respondents' demographic characteristics, including household income, educational attainment, Internet use, and prior experiences choosing schools.

Table 1. Demographic characteristics of the analytic sample

	Number	Percentage
Parent characteristics		
Female	2,642	75.5
Race/ethnicity		
White, non-Hispanic	2,482	70.9
Black, non-Hispanic	422	12.1
Hispanic	358	10.2
Other, non-Hispanic	238	6.8
Age		
Younger than 25	233	6.7
25–34	1,276	36.5
35–44	1,055	30.1
45 or older	936	26.7
Education		
Less than high school	113	3.2
High school diploma or GED	1,955	55.9
Associate's degree	776	22.2
Bachelor's degree or higher	656	18.7
Language spoken at home		
English only	3,254	93.0
Spanish	185	5.3
Other	61	1.7
Marital status		
Married or living with a partner	2,293	65.5
Respondent's role in child's education decisions		
Only person who makes decisions	1,214	34.7
The main person, but takes into account the opinion of the child or another adult	1,058	30.2
Share equally in the decision with the child or other adult	1,133	32.4
Involved, but in some other way	95	2.7
Household income, annual		
$10,000 or less	351	10.0
$10,001–$20,000	696	19.9
$20,001–$30,000	1,194	34.1
$30,001–$40,000	1,259	36.0
More than $40,000	0	0.0
Internet use per week		
Less than 10 hours	786	22.5
10–29 hours	1,712	48.9
30 or more hours	974	27.8
Characteristics of parent's youngest child		
Child is female	1,677	47.9
Child has ever had an individualized education plan	805	23.0
Community characteristics		
Public school options other than the neighborhood school available in the community (e.g., magnet, charter)	1,819	52.0
Urbanicity		
Urban	1,114	31.8
Suburban	1,307	37.3
Rural	1,079	30.8

Source: Parent information and school choice survey administered in August–October 2016.

Note: These demographics are for the analytic sample of 3,500 parents.

Participants took part in the experiment by logging into a website and completing a series of steps as follows. First, they completed a baseline survey. Next, each study participant (parent) was randomly assigned with equal probability to one of the study's 72 information displays. Each participant saw only one information display (in a separate window of a web browser), showing profiles of 16 elementary schools from a hypothetical district. Each of the 72 displays represented data in different ways but described the same 16 schools. We designed these schools' profiles to differ from one another with respect to their distance from the chooser's hypothetical home and the information reported about school academics, safety, and resources. We generated this hypothetical information to create trade-offs between schools—for example, the school located closest to home also had relatively poor academic performance, low safety, and few resources compared with other schools. We reported these features as the following:

- Academic performance: proficiency rates based on state assessments
- Safety: percentage of students who had never received a suspension (that is, higher numbers indicate safer schools)
- School resources: number of laptops or tablets per 100 students

The third part of the study was an endline survey that appeared after parents had reviewed the school profiles in their assigned information display. We used the endline survey to measure the study's three outcomes. First, we asked parents to identify and rank their three favorite schools from the full set of 16. The purpose of this ranking task was to learn how information displays affected parents' behavior, in terms of how they choose and rank schools. To measure choice outcomes, we examined the attributes of top-ranked schools. As mentioned above, the report cards displayed four categories of school attributes: distance, academics, safety, and resources.

Next, we collected outcome data related to parents' factual understanding of the information presented. The school information display remained visible during the endline survey, so parents could refer to it while responding to these items and all other items in the endline survey. The understanding outcome was the percentage of nine items that parents answered correctly.

Finally, the survey included attitudinal questions to gauge parents' impressions of the displays' ease of use, and their satisfaction with the displays. Each question had a four-point scale, asking respondents to agree strongly, agree, disagree, or disagree strongly with a particular statement related to satisfaction or ease of use. We used four survey items to measure satisfaction and five items to measure ease of use. For each measure, we calculated the percentage of items with which respondents agreed or agreed strongly.[2]

Summary descriptive statistics for each of these outcomes, including the characteristics of the schools that were used in the study (to be ranked by parents) and the knowledge and attitude scales are shown in Table 2.

[2] Details on the scales, including question wording, item correlations, and factor analysis results are available in the online appendix.

Table 2. Descriptive statistics for outcome measures

Outcome measure	Scale	Description	Mean	Standard deviation
Choice outcome: Characteristics of schools (N = 16)				
School academics	0–100	Percentage of students proficient on state standardized test	65.0	23.0
School distance from home	0–6	Miles	2.0	1.4
School safety	0–100	Percentage of students with no suspensions	92	5
School resources	0–100	Number of laptops or tablets per 100 students	55	20
Knowledge and attitude outcomes (N = 3,500)				
Understanding	0–100	Percent correct (of 9 items)	73	30
Ease of use	0–100	Percentage of statements agreed or agreed strongly	86	23
Satisfaction	0–100	Percentage of statements agreed or agreed strongly	90	21

The experiment's implementation was successful. Overall attrition (5.5%) and differential attrition rates by factor level (which ranged from 4.8% to 5.8% across factor levels) in the experiment were low. Baseline equivalence tests for the final analytical sample confirmed that the study's randomization procedures produced equivalent samples across each of the tested factor-levels in the study (see the online analytical appendix for additional details).

B. Analysis methods: Hierarchical Bayesian analysis

The analysis relied on Bayesian methods to estimate the effects of each factor level. Historically, factorial experiments in education research have been rare, in part because of the large sample sizes required. In a conventional analysis of results from a factorial design, each factor level tested in an experiment would require its own independent hypothesis test. As the number of hypothesis tests increases, so does the probability that at least one of them will yield a false positive—a situation referred to as the multiple comparisons problem (Waller & Duncan, 1969). Although correcting for multiple comparisons is possible, the most common ways of doing so effectively apply a post-hoc penalty on the precision of the experiment. As a result, acquiring a sample size large enough to test more than a few treatment arms is often difficult. In contrast, a Bayesian model specifies the joint probability distribution for all the parameters in the model using a prior distribution. This neutralizes concerns about Type I errors, which are incorrect rejections of the null hypothesis, because we no longer must maintain a null hypothesis that the true effect equals zero (Gelman, Hill, & Yajima, 2012).

Hierarchical Bayesian analyses also incorporate information about the expected relationships between parameters of interest, which has important consequences for a factorial experiment, with two chief benefits. The first is to achieve "partial pooling," which can improve the statistical precision of effect size estimates profoundly (Gelman, 2005; Kassler et al., 2018). The term partial pooling refers to the process of pooling observations across all levels of a factor when estimating the effect of each of the factor's levels, especially when that factor has little effect on outcomes compared with other factors. Within a given factor, the result is that the estimates for the effects of each level are informed by one another, leading to larger effective sample sizes and smaller uncertainty in estimates. The variance parameters of these priors are also partially pooled to borrow information about the overall effect size across factors, providing greater stability in estimates of factors with few levels (Gelman & Hill, 2007).[3]

The experiment defined treatment arms with a set of five factors, described previously. In addition, we made the model scale free; that is, all outcomes were standardized to have a mean of zero and a standard deviation of 1, as were all continuous predictors; binary predictors were left as 0/1. The study analyzed data from respondents in all 72 treatment arms to estimate the following model:

$$y_i = \alpha + \sum_{m \in F} \beta^{(m)}_{j_i^{(m)}} + \sum_{\substack{q,r \in F \\ q \neq r}} \theta^{(q,r)}_{j_i^{(q)}, j_i^{(r)}} + \gamma \cdot X_i + \varepsilon_i.$$

In the equation above, respondents are indexed by i, so that y_i is the outcome of interest for respondent i. The set F is a set of indices representing the five factors in the experiment. For a given factor $m \in F$, the index $j_i^{(m)}$ indicates the level of factor m respondent i receives. The term $\beta_j^{(m)}$ represents the main effect of factor m at level j, and the term $\theta_{k,l}^{(q,r)}$ represents the interaction effect between factor q at level k and factor r at level l. Thus, the term $\beta^{(m)}_{j_i^{(m)}}$ in the equation above represents the main effect of factor level j of factor m on the outcome of respondent i.The vector X_i is a set of additional covariates with effects γ, α is an overall intercept, and ε_i is a respondent-level error term.[4]

The prior distributions for the model's parameters are as follows:

$$\begin{aligned}
&\beta^{(m)} \sim \mathcal{N}\left(0, \tau^{(m)}\right) \\
&\theta^{(q,r)} \sim \mathcal{N}\left(0, \tau^{(q,r)}\right) \\
&\epsilon \sim \mathcal{N}(0, \sigma)
\end{aligned}$$

$$\begin{aligned}
&\tau^{(m)} \sim \mathcal{N}(0, \phi_{\text{main}}) \\
&\tau^{(q,r)} \sim \mathcal{N}(0, \phi_{\text{int}}) \\
&\alpha, \sigma, \gamma, \phi_{\text{int}}, \phi_{\text{main}} \sim \mathcal{N}(0,1)
\end{aligned}$$

Here $\mathcal{N}(0, s)$ indicates either a normal distribution with a mean of zero and standard deviation s, or the corresponding half normal for the standard deviation parameters τ, σ, and ϕ

[3] These within-factor variance components can be interpreted as a gestalt measure of the importance of each factor on the outcome. These variance parameters are themselves modeled as coming from a common prior that reflects expectations about the overall distribution of effect sizes in the experiment. In the parlance of Bayesian statistics, the parameters of the prior distribution are known as *hyperparameters*, and the priors on the hyperparameters as *hyperpriors*. An astute statistician will note that one could model the parameters of the hyperpriors with priors of their own, and so on. This is unnecessary in practice, and the aspiring Bayesian statistician need not worry about continuing to define even higher-level parameters to govern these hyperpriors. Although partial pooling is not a uniquely Bayesian approach—non-Bayesian mixed models can achieve a similar effect—estimating the variance components for factors with a very small number of levels in a non-Bayesian setting would not be possible. Using the hyperprior on the variance components allows us to do this (Gelman & Hill, 2007, pp. 498-500).

[4] We ruled out third-order and higher-order interaction effects. The small size of second-order effects that we report in the findings suggests this assumption is reasonable.

(the term *half normal* refers to a normal distribution truncated below a value of zero, meaning there are no negative values). The first three rows here define priors for the parameters of main interest in the model, while the next two rows define priors for the parameters of these priors. The last row sets the prior for parameters we do not want to model with additional structure or strong prior information, using a distribution that is broad and relatively uninformative on the scale of the model. In selecting these priors, we followed previous work (Gelman, 2006) and the current recommendations from the Stan Development Team (2017).

Rather than pick a baseline or reference level for each factor in the model, we explicitly include a term $\beta_j^{(m)}$ for every level of each factor in our model. To preserve identifiability of the model, we impose the constraint that the main effects for the levels of each factor must sum to zero: $\sum_m \beta_m^{(m)} = 0$. The effect of a factor is read off relative to zero (and zero is by definition the mean of the effects for each factor). We also prefer this approach for the sake of interpreting our results, as no clear baseline category exists for the school information design strategies tested in our experiment—that is, for every factor, a decision must be made. We use analogous contrasts for the interaction terms: we explicitly model an interaction term $\theta_{p,q}^{(p,q)}$ for each combination of levels of each pair of factors and impose the constraint that these effects sum to zero within each pair of factors: $\sum_{p,q} \theta_{p,q}^{(p,q)} = 0$. This choice of contrasts for interaction effects means the expected effect of a given factor level is not, in general, equivalent to the main effect $\beta_m^{(m)}$ of that level read directly from the model. To read off the full effect of a given factor level, we add to the main effect the average of all interaction terms that involve that factor level: the total effect of factor m at level $j^{(m)}$ is given by $\beta_{j^{(m)}}^{(m)} + \sum_{\substack{q \in F \\ q \neq m}} \frac{1}{J^{(q)}} \sum_{j^{(q)}} \theta_{j^{(m)},j^{(q)}}^{(m,q)}$, where $J^{(q)}$ is the number of levels of factor q.

IV. FINDINGS

A. Impacts on choices: nudge effects

1. Individual factor levels

The impact of different design decisions on choice are shown in Table 3. Evidence suggests that even subtle changes can nudge parents toward choosing one type of school or another simply by varying the presentation of information. Changing the default from sort by distance to sort by academics had the strongest effect, leading choosers to select a school that was both higher performing academically (a difference of 0.20 standard deviations in proficiency rates, equivalent to 5 percentile points in the school ranking), and farther away (a difference of 0.38 standard deviations in distance, equivalent to 0.6 miles). These substantial effects were particularly noteworthy because all of the choosers in the experiment had the option of re-sorting schools according to either of these criteria.

Table 3. Impacts on characteristics of selected schools (nudge effects)

Factor	Factor level	Effect size for: Academics	Distance	Resources	Safety
Format	Numbers only	-0.05	0.01	0.02	-0.02
	Numbers + icons	**0.05***	-0.03	-0.04	**0.12***
	Numbers + graphs	0.00	0.02	0.02	-0.11
District average shown	No	-0.02	0.02	0.01	-0.03
	Yes	0.02	-0.02	-0.01	**0.03***
Source of information	District only	-0.01	-0.01	**0.06***	-0.06
	District + parent ratings	0.01	0.01	-0.06	**0.06***
Amount of information	Lower amount	**0.04***	0.00	-0.09	0.00
	Progressive disclosure	**0.05***	-0.03	-0.09	0.01
	Higher amount	-0.09	0.02	**0.18***	-0.01
Default sort order	By distance	-0.10	**0.19***	-0.01	-0.01
	By academics	**0.10***	-0.19	0.01	0.01

Notes: Each column of the table summarizes the results of a separate regression. The effect sizes represent, in standard deviation units, the effect of each display strategy on the average z-score of selected schools within a given category of information. The asterisk (*) and bold blue text indicate when a strategy is likely to have impacted selections: effect sizes are highlighted when the probability of the strategy having a true effect greater than zero was greater than 0.70 (calculated from a Bayesian posterior distribution).

Displaying information about academics and safety using icons (while omitting icons for distance and resources) also had an influence on choice. When icons were added to the display, compared with a display showing numbers only, choosers shifted their school rankings to favor schools with higher scores on academics (difference of 0.10 standard deviations) and safety (0.14 standard deviations)—the focus of the letter grades—and away from distance (–0.04 standard deviations) and resources (–0.06 standard deviations).

The amount of information was influential as well. When the level of detail shifted from the lower-information condition or the progressive disclosure condition toward the higher-information condition, choosers shifted toward ranking schools based on resources (difference of 0.27 standard deviations), the attribute for which the most new information was revealed in the high-information condition.[5]

Another nudge effect that emerged from the experiment was related to the source factor, specifically, whether the display included parent survey information to describe schools' safety and academics. We found that when the display included parent survey results, parents chose schools with higher ratings on the safety dimension (difference of 0.12 standard deviations) and lower ratings on the resources dimension (-0.12 standard deviations).

[5] We assigned characteristics to the hypothetical schools in such a way that the rank ordering on any attribute was unchanged whether one saw the high-information or low-information display. We added enough jitter to the values so this relationship would not be obvious to those who could see both the summary and the detail.

We also examined two-way interaction effects, whether changing any pair of factor levels at once had a greater impact than the sum of the individual factor-level effects. We did not find meaningful impacts on choices beyond the main effects discussed so far.

2. Cumulative effects

The design choices that we made were independent from one another, meaning that they could be combined to additively produce a desired outcome. In Table 4 we present these impacts in terms of percentiles rather than effect sizes, to describe the change in the relative rankings of the schools parents selected after viewing various displays. For example, the display that was the most successful at nudging parents toward choosing a school based on academic performance sorted by academic performance by default; displayed a low amount of information; included parent survey ratings and icons for academics and safety; and included a district reference. This display strategy resulted in parents choosing schools with academic performance 19 percentile points higher than the performance of schools they might have chosen had they been presented with the worst combination (sorting by distance, with a high amount of information, no parent surveys, no icons, and no district reference). We also observed cumulative effects for nudges toward choosing schools based on distance from home, safety, and resources. The difference between the best and worst display strategy for each type of nudge ranged from 17 percentile points for the safety nudge to 21 percentile points for the distance nudge.

Table 4. Cumulative effects: Predicted effects on choice for best and worst factor combinations

	Factor combination						
Outcome	Source: Includes parent survey	Default sort	District reference	Format	Amount of information	Predicted mean	Best–worst
Academic performance of chosen schools (percentile)							
Best display	Yes	Academics	Yes	Icons	Low	58.7	19.3
Worst display	No	Distance	No	Numbers	High	39.4	
Proximity of chosen schools (percentile)							
Best display	Yes	Distance	No	Graphs	High	60.3	20.9
Worst display	No	Academics	Yes	Icons	Progressive disclosure	39.4	
Safety of chosen schools (percentile)							
Best display	Yes	Academics	Yes	Icons	Progressive disclosure	59.1	17.0
Worst display	No	Distance	No	Graphs	High	42.1	
Resources of chosen schools (percentile)							
Best display	No	Academics	No	Numbers	High	60.6	18.6
Worst display	Yes	Distance	Yes	Icons	Low	42.1	

Note: A percentile score of 50 represents the school with the median value of the attribute among the 16 schools in the hypothetical district used for the study.

B. Impacts on understanding, satisfaction, and ease of use

1. Individual factor levels

Only three of the five factors—**format**, **source**, and **information amount**—on their own had a substantial impact (effect size difference of at least 0.05) on satisfaction, understanding, or ease of use. Table 5 shows the results. For the format factor, parents found the numbers-only display easier to understand than icons or graphs (effect size of 0.06 versus –0.03). This difference of 0.09 standard deviations is based on shifting the outcome score from 72% to 75% correct. The probability that the numbers-only strategy outperforms the others is 99%. We also found that including parent ratings of schools increased the predicted level of satisfaction from 89% to 92%, inducing an effect size difference of 0.12 standard deviations. And displaying a higher amount of information increased the predicted level of satisfaction from 89% to 91% (effect size difference of 0.06) compared to both the lower-information display and progressive disclosure.

Table 5. Impacts on understanding, ease of use, and satisfaction

		Posterior mean[a]			Effect size[b]			Probability of being the best strategy[c]		
Factor	Strategy	Understanding	Ease of use	Satisfaction	Understanding	Ease of use	Satisfaction	Understanding	Ease of use	Satisfaction
Format	Numbers only	**75.0**	86.0	90.0	**0.06**	0.00	0.00	**0.99**	0.23	0.26
	Numbers + icons	72.5	86.0	89.5	-0.03	0.00	-0.03	0.01	0.43	0.03
	Numbers + graphs	72.0	86.0	**90.5**	-0.03	0.00	**0.03**	0.00	0.34	**0.71**
District average shown	No	73.5	**86.5**	**90.5**	0.00	**0.01**	**0.02**	0.52	**0.84**	**0.86**
	Yes	73.5	86.0	89.5	0.00	-0.01	-0.02	0.48	0.17	0.14
Source of information	District only	**73.5**	86.0	89.0	**0.01**	0.01	-0.06	**0.84**	0.68	0.00
	District + parent ratings	73.0	86.0	**91.5**	-0.01	-0.01	**0.06**	0.16	0.32	**1.00**
Amount of information	Lower amount	73.5	86.5	89.5	0.00	0.02	-0.02	0.32	0.66	0.04
	Progressive disclosure	73.5	86.0	89.5	0.00	0.00	-0.02	0.42	0.23	0.06
	Higher amount	73.0	86.0	**91.0**	0.00	-0.01	**0.04**	0.26	0.11	**0.91**
Default sort order	By distance	**74.0**	86.0	**90.5**	**0.02**	0.00	**0.02**	**0.89**	0.50	**0.87**
	By academics	73.0	86.0	89.5	-0.02	0.00	-0.02	0.11	0.50	0.13

Note: Posterior probabilities above 0.70, and the corresponding effect sizes and posterior means, are indicated with bold blue text.

[a] Posterior means are the average predicted score for the full sample, holding the factor constant at the given factor level.

[b] These columns report the impact of each display strategy on the relevant outcome in effect size (standard deviation) units, with effects estimated relative to the average outcome across all combinations of strategies in the experiment.

[c] The probability of being best, derived from the Bayesian posterior, is the probability that a given strategy outperforms the other tested strategies for that factor. Within a factor, these estimated probabilities always sum to 1 for a given outcome.

Although the estimated impacts may have been small for most factors and outcomes, most of the comparisons still yielded a clear winner, meaning that the probability that one factor level produced a better outcome than the other(s) was greater than 70%. Of the 15 sets of comparisons (five factors x three outcomes), nine had a factor level whose posterior probability of being best was 70% or greater, the arbitrary threshold we chose to determine that a decision was actionable. The prior probabilities of any factor being best were 50%–50% for factors with two levels, and 33%–33%–33% for factors with three levels.

Even if the absolute size of the differences in predicted effects was small, the information is actionable. First, the effects are cumulative, and several decisions with small effects can add up to a substantial effect, as discussed below. Second, designers of school choice information displays must choose among alternatives that are generally cost-neutral, such as whether to display or suppress information that is already available, or which option to set as the default. The choice can be intentional or not, but there is no avoiding the choice. Knowing with a high degree of certainty that one is better than the other, even if the difference is as small as 0.02 standard deviations (as is the case with the smallest effects in this study that have a high posterior probability of being greater than zero), could be sufficient to choose that factor level. Exceptions might be if the factor level cannot be costlessly implemented. For example, parent survey data might not exist in certain cases, so the source-of-information results would have to be considered in the context of the costs of obtaining those data.

The finding that the effects on understanding, satisfaction, and ease of use were small does not suggest that design, in general, does not matter. The particular factors and factor levels were implemented in this study by a design firm using established best practices. We purposely avoided including “straw man” factors or factor levels that we knew would be dominated. Thus, the study findings should not be seen as evidence generally that school report card design does not affect users’ ability to understand or interact effectively with the school data. Plainly bad designs would likely have negative consequences.

2. Cumulative effects

Although the effects of each factor level on its own may be small, the posterior means of the cumulative effects of changing several factors at once can be quite meaningful. We might examine impacts when varying one factor at a time, but the decision maker must decide on all five (and many more) simultaneously. Table 6 shows that even among the “best practices” options tested in this experiment, the predicted understanding score can vary by more than 5 points, from 71% for the worst (understanding-minimizing) factor combination to 76% for the best (understanding-maximizing) factor combination. Ease of use did not vary more than 1 percentage point, but satisfaction varied by more than 6 points, from 87% to 93%. Figures 2–4 show the factor combinations that produce the best predicted results for each outcome.

Table 6. Cumulative effects: Predicted knowledge and attitudes for best and worst factor combinations

Outcome	Factor combination					Predicted mean	Best–worst
	Source: Includes parent survey	Default sort	District reference	Format	Amount of information		
Understanding (mean on 0–100 scale)[a]							
Best display	No	Distance	No	Numbers	Progressive Disclosure	76.3	5.1
Worst display	Yes	Academics	Yes	Graphs	High	71.1	
Ease of use (mean on 0–100 scale)[b]							
Best display	No	Academics	No	Icons	Low	86.7	1.2
Worst display	Yes	Distance	Yes	Numbers	High	85.5	
Satisfaction (mean on 0–100 scale)[c]							
Best display	No	Distance	No	Graphs	High	93.3	6.2
Worst display	Yes	Academics	Yes	Icons	Low	87.1	

[a] The understanding score is the percentage of factual questions about school attributes that respondents answered correctly.

[b] The ease-of-use score is the percentage of statements about the information being easy to use with which respondents agreed or agreed strongly.

[c] The satisfaction score is the percentage of statements related to satisfaction with which respondents agreed or agreed strongly.

Figure 2. Best display for understanding

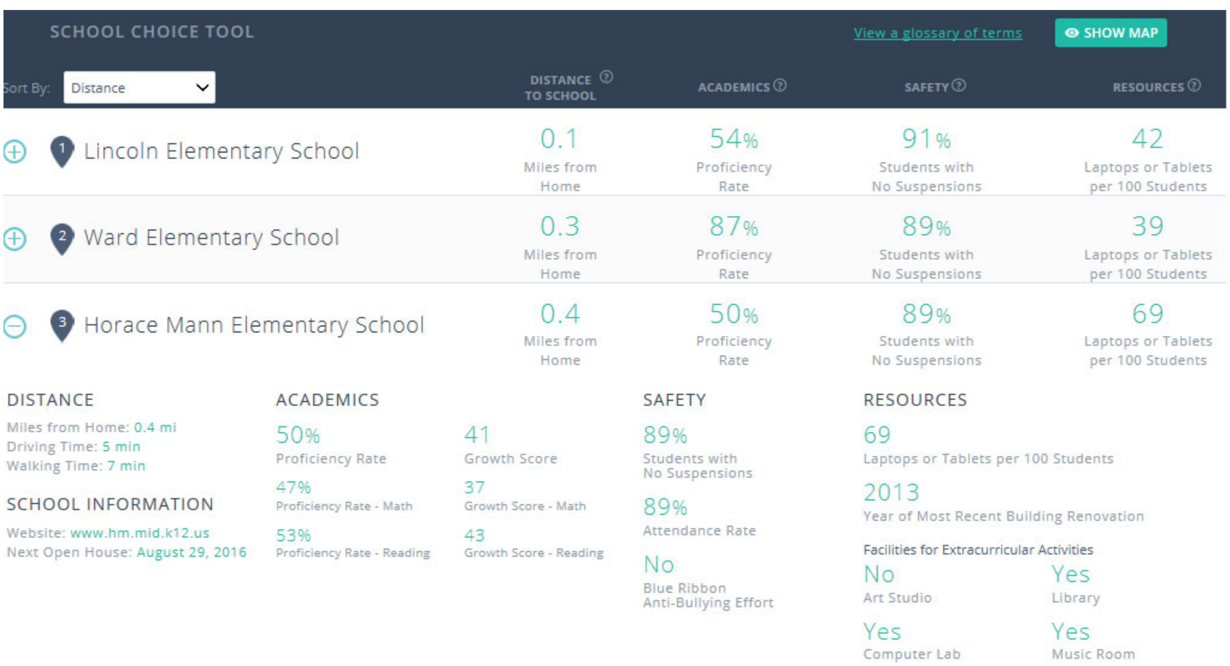


Figure 3. Best display for ease of use

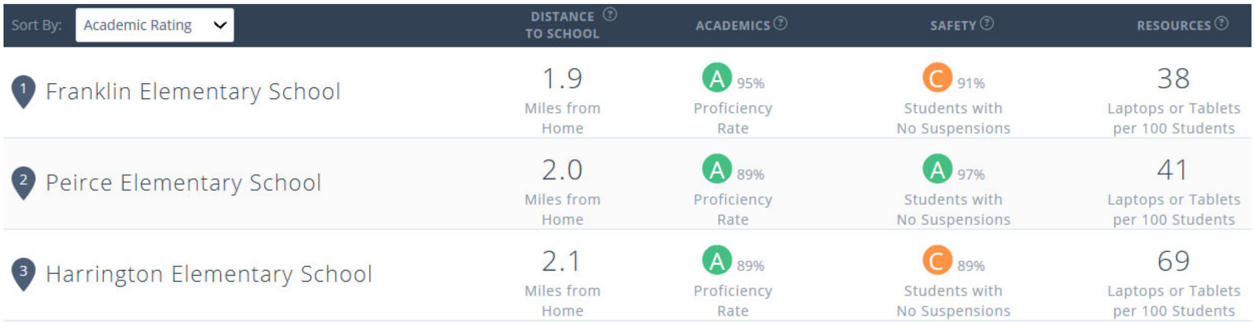
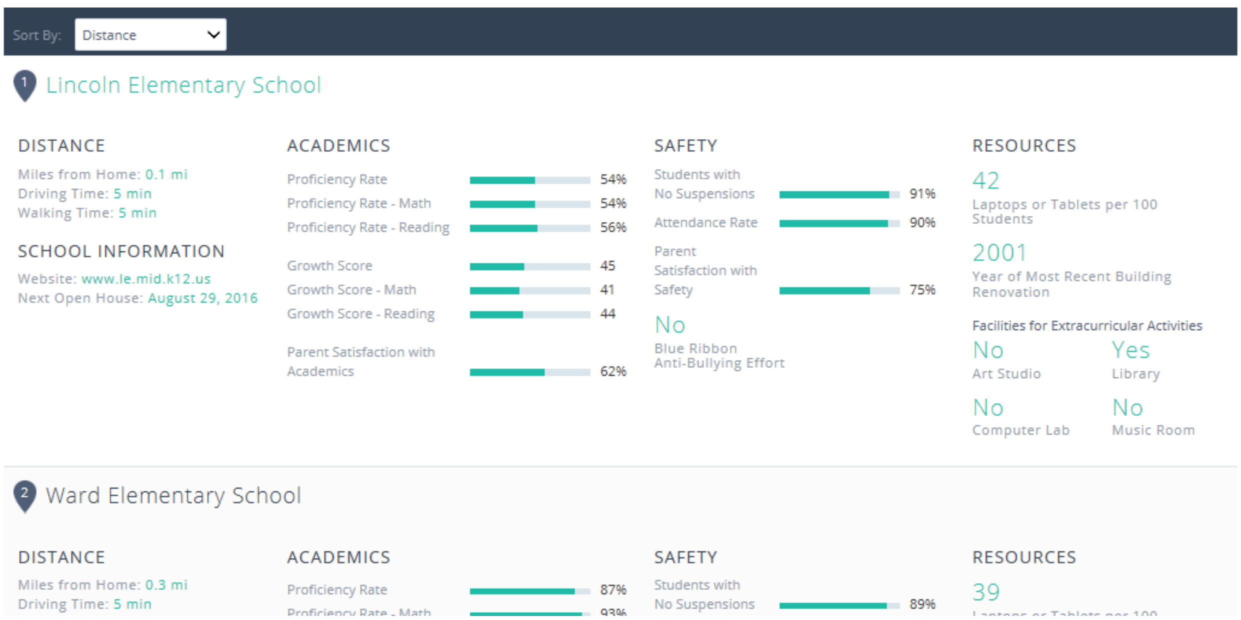


Figure 4. Best display for satisfaction



3. Trade-offs

In terms of knowledge and attitudes, more than half of the factor-outcome combinations produced a clear winner, but there is a trade-off to consider for three of those factors. Which design choice is appropriate depends on which outcome one values most. For example, we noted above that the numbers-only format increased understanding, but it also reduced satisfaction. Conversely, including parent ratings in the display increased satisfaction, but *reduced* the level of understanding by a small amount (Table 5). For the other three factors, the preferred factor level was unambiguous. Including district averages as a reference point reduced ease of use and satisfaction, so it is not recommended. The more detailed display dominated both the less detailed display and even the version that allowed users to click through to a more detailed display (progressive disclosure). Finally, sorting by distance as a default dominated the option to sort by academic performance as a default, with small improvements in both understanding and satisfaction and no penalty to ease of use.

To the extent that information architects care about nudging consumers toward particular choice outcomes rather than improving their understanding, satisfaction, and user experience with the information display during the choice process, there are still other trade-offs to consider. Sorting the data by academic performance does lead consumers to choose on that basis, but it also leaves them slightly less satisfied and less able to understand the data compared with a distance-sort default.

C. Robustness of findings

The results presented above were robust to including interaction effects, focusing on subsamples, and using alternative model specifications. Across all of the outcomes in the study, the analysis showed that interaction effects did not meaningfully change the core pattern of results summarized by the main effects of each display strategy. The study's analytical model calculated pairwise interaction effects between factor levels. (For example, the model estimated whether the effect of displaying a larger amount of information on understandability became larger or smaller when the display included graphs.) In practice, these interaction effects did not prove to be important: the interaction effects ranged in size from –0.03 standard deviations to 0.02 standard deviations, and 91% of the calculated interaction effects fell within -0.01 to 0.01 standard deviations. None of the interaction effects was large enough to change conclusions about which factor level was best for each outcome (based on the study's 70% threshold).

The results are also largely robust to trimming the sample in ways that may be of particular interest to policymakers and practitioners. The sample for the study was defined to encompass low-income parents because this was a way to represent the target population of disadvantaged choosers. As Table 1 shows, the study population consists of parents of school-age children with less than $40,000 in annual income. However, the low income threshold might still include some families who might not be considered disadvantaged. Therefore, we reestimated the results on a subsample for which we applied a lower income threshold ($30,000 instead of $40,000 per year) and another subsample for which we excluded anyone with a college degree to rule out high-education sample members whose low-income status might be temporary. We also reestimated the results for a subsample defined as being less Internet savvy (use the Internet for less than 30 hours per week) and anther subsample defined as having reported experience with school choice.

Most of the study-estimated effects remained largely consistent if the sample is restricted to each of these subsamples (Table 7). Of the 18 separate effect sizes the study estimated, 16 remained consistent for parents with lower incomes, 13 remained consistent for parents with lower education, 17 remained consistent for parents with non-intensive Internet use, and 15 remained consistent for parents with experience with school choice. However, it is important to note that the sample sizes are smaller in each of these subsamples, so the subsample results are estimated with less precision than the full sample.

Table 7. Comparison of subsample results to results for the full sample

Research question	Outcome	Best strategy (largest impact, posterior probability >70%)				
			Subsample			
		Full sample (N = 3,500)	Prior school choice experience (N = 1,819)	No college completion (N = 2,068)	Income < $30,000 per year (N = 2,241)	Internet use < 30 hours per week (N = 2,498)
1. Format	Understanding	Numbers only	*C*	*C*	*C*	*C*
	Ease of use	No differences	*C*	*C*	*C*	*C*
	Satisfaction	Numbers + graphs	*Numbers + graphs not best strategy*[a]	*Numbers + graphs not best strategy*[a]	*Numbers + graphs not best strategy*[a]	*C*
2. District average shown	Understanding	No differences	*Without district average is best*	*Without district average is best*	*C*	*C*
	Ease of use	Without district average	*C*	*C*	*C*	*C*
	Satisfaction	Without district average	*C*	*C*	*C*	*C*
3. Source of information	Understanding	District only	*C*	*C*	*C*	*C*
	Ease of use	No differences	*District only is best*	*District only is best*	*C*	*C*
	Satisfaction	District + parent ratings	*C*	*C*	*C*	*C*
4. Amount of information	Understanding	No differences	*C*	*C*	*Progressive disclosure is best*	*C*
	Ease of use	No differences	*C*	*C*	*C*	*C*
	Satisfaction	Higher amount	*C*	*C*	*C*	*C*
5. Default sort order	Understanding	By distance	*C*	*C*	*C*	*C*
	Ease of use	No differences	*C*	*By academics is best*	*C*	*By academics is best*
	Satisfaction	By distance	*C*	*By academics is best*	*C*	*C*
6. Effects on choices	Encouraging academics	Sort by academics	*C*	*C*	*C*	*C*
	Encouraging academics	Format with numbers + icons	*C*	*C*	*C*	*C*
	Encouraging academics	Lower amount of information or progressive disclosure	*C*	*C*	*C*	*C*

[a] The format with the largest effect size for this outcome in this subsample (numbers only) differed from the overall sample. The probability that the numbers-only format was the best did not exceed 70%, but the subsample finding is inconsistent with the full sample finding by the "different best strategy" rule discussed in Section B.1.
C = subsample finding is consistent with full sample finding.

We also confirmed that the results are robust to two plausible alternative model specifications. First, we estimated effects without including any demographic covariates for the survey respondents (omitting controls for household income, parent's education, computer use, experience with school choice, and the youngest child's gender, race/ethnicity, and special education status). Second, we tested whether any evidence indicates that a frequentist analysis (that is, a model using a standard frequentist regression framework with non-informative, non-hierarchical priors) would have arrived at different conclusions. Neither sensitivity test produced results that were meaningfully different from the main model: the magnitude and direction of effect sizes were consistent across all three tests. For the understanding outcome, the absolute differences between the original and any of the alternatives ranged from 0.00 to 0.02 standard deviations (Table 8). However, for the frequentist analogue sensitivity test, using a conventional significance test with a *p*-value of 0.05 did cause a substantial loss of precision compared with the Bayesian model. Only one of the effects (the effect of the numbers-only format on understanding) was statistically significant under this approach, even though all other point estimates remained similar in magnitude and sign to the estimates in the study's preferred Bayesian model.

Table 8. Sensitivity tests for the understanding outcome

		Original	No covariates	Frequentist analogue
Format	Numbers only	0.06	0.07	0.08*
	Numbers + icons	-0.03	-0.03	-0.03
	Numbers + graphs	-0.03	-0.04	-0.05
District average shown	No	0.00	0.00	0.00
	Yes	0.00	0.00	0.00
Source of information	District only	0.01	0.01	0.02
	District + parent ratings	-0.01	-0.01	-0.02
Amount of information	Lower amount	0.00	0.00	0.00
	Progressive disclosure	0.00	0.00	0.01
	Higher amount	0.00	0.00	-0.01
Default sort order	By distance	0.02	0.02	0.02
	By academics	-0.02	-0.02	-0.02

Note: The table reports the impact of each display strategy on the relevant outcome in effect size (standard deviation) units, with effects estimated relative to the average outcome across all combinations of strategies in the experiment.

* For the frequentist analogue sensitivity test, results were statistically significant at the 0.05 level with a Benjamini-Hochberg adjustment for multiple comparisons that accounts for the total number of display strategies (12) in the experiment.

V. CONCLUSIONS AND POLICY IMPLICATIONS

This experiment shows that subtle differences in the ways information is presented to consumers, such as default settings, can change their behavior in important ways. The default sort order in a list of schools, for example, is trivial to change in an online display but influenced school choice by a notable margin. This is consistent with much prior research showing that defaults matter (Johnson & Goldstein, 2003). Importantly, this is a choice that designers cannot avoid, even if users have the option to re-sort the data.

Defaults also played a role in the study's progressive disclosure condition. Designers choosing between higher and lower density information displays have the option of allowing users to choose the level of detail they see, but designers still must decide whether the default is low information with an option to expand detail, or high information with an option to hide detail. This study defaulted to low information and, indeed, many of the progressive disclosure outcomes looked like those of the low-information condition, suggesting that parents might have used the tool with the defaults in place, not expanding the display to reveal the detail. With the platform used in the current experiment, tracking the number of times the progressive disclosure option was exercised was not possible, but future experiments of this type could strive to measure user interaction with the site in this way.

We also found evidence of nudges from the way school information is formatted and the amount of information presented. Adding icons to the display led parents to choose higher-performing schools. Lower information also led to selection of higher-performing schools. These nudge factors have one characteristic in common: making an option or piece of information more prominent increases the likelihood that users will prioritize that information.

The experiment showed only modest effects on consumers' understanding of the choices and their user experience. The key insight here is that the impacts themselves, although modest, sometimes worked in opposite directions. For instance, one set of choices can maximize understanding, but those decisions may come at the expense of ease of use. Similarly, the evidence suggests there are trade-offs between satisfaction and ease of use, or that an approach that is both satisfying and easy to use (default sort by distance instead of academic performance) might produce less socially desirable outcomes, such as reducing focus on academics, and hence reduce the incentives for schools to maximize academic performance when trying to attract parents. When designers make information easier to digest, one consequence is that the information might not be as satisfying and consumers might not be as well informed.

There are some important limitations to keep in mind when interpreting these results. The first has to do with external validity. The experimental estimates, by using random assignment without any danger of contamination or crossover, were internally valid causal estimates of the impact of information displays on the measured outcomes. However, the study was an online experiment with fictional schools. In field settings, consumers already have information about schools through word of mouth, school visits, schools' own websites and advertising, and personal experiences with schools (Schneider, Teske, & Marschall, 2000). The effects of online information profiles could be more modest in this context.

A related difference between this experiment and actual school choice settings is that the stakes are much higher in real life, so the nudge effects of online displays might be smaller than what we observed in the experiment. This is because consumers will invest more in the choice, perhaps being less influenced by default settings (e.g., intentionally re-sorting the data or expanding a listing if those options exist). Still, it is important to note that nudges do not exist only at the time of choosing. It could be, for example, that the sort order of a list that parents saw months before submitting an application affected which schools they decided to tour or research on the Internet—ultimately affecting their choices, albeit more circuitously.

Another limitation of this study is that the mechanisms by which the nudges worked are unclear in some cases. For example, the icons tended to highlight differences in academic performance by grouping schools into a small number of (three) different categories of performance, and the pattern of effects may be different when icons map onto a larger number of categories. Similarly, the low-information condition gave equal space to all four domains of school information, but the high-information condition displayed disproportionately more information about school resources. In this way, the "amount" factor could be a prominence effect, reflecting the percentage of screen space dedicated to the information domain rather than the amount.

This study has implications for research. In particular, Bayesian hierarchical models offer a promising strategy to test the effects of several experimental manipulations at once. Researchers who find themselves interested in testing more than one factor in an experiment but limited by concerns of statistical power may find substantially more opportunity with a Bayesian factorial design. In this study, we manipulated five factors at once, with precise estimates for each main effect and two-way interaction effects. This would not have been possible with a conventional factorial experiment and the multiple comparisons corrections it demands. A Bayesian approach could enable researchers to pursue factorial experiments in contexts in which they otherwise would not seem possible.

This study also has implications for policy. The past few decades have produced education reforms in the United States that ask parents to make choices for their children. These reforms have made parents' decision making extremely consequential, both for their own children (in determining which schools they attend) and the broader functioning of U.S. schools (in determining what types of schools flourish, and what types fail, in school choice markets). Knowingly or not, policymakers affect these choices. Even seemingly mundane decisions about the order in which schools appear and whether data are presented graphically nudge parents toward one type of school or another. As we have argued, these nudges are unavoidable. For example, listing schools alphabetically might seem neutral and nudge-free, but researchers have found that alphabetically listing candidates on election ballots (Miller & Krosnick, 1998) and student applicants to selective universities (Jurajda & Münich, 2010) benefits those with names that appear earlier in the alphabet. Given what our results show about the substantial impacts for the default sort ordering of schools, the same could very well be true in the school choice context.

This study provides evidence on a number of specific questions about the presentation of school choice information. For example, it shows that default sort order affects parents' choices and that using graphs to represent school performance data can introduce trade-offs between satisfaction and understanding. However, we believe the most important implication might be broader than that: that choice architecture matters in education. Policymakers shape parents' choices through the decisions they make, often in subtle ways. Perhaps some decisions—such as allowing users to customize the information they see—could soften the intensity of these nudges, but nudges are an inescapable reality of providing information. Given this reality, decisions about how to present information about schools should be made carefully, with thoughtful attention to the nudges that might result—and the consequences for students and schools.

REFERENCES

Bar-Hillel, M. (2015). Position effects in choice from simultaneous displays: A conundrum solved. *Perspectives on Psychological Science*, *10,* 419–433.

Bell, C. A. (2009). All choices created equal? The role of choice sets in the selection of schools. *Peabody Journal of Education,* 84(2), 191–208.

Borgmeier, I., & Westenhoefer, J. (2009). Impact of different food label formats on healthiness evaluation and food choice of consumers: A randomized-controlled study. *BMC Public Health, 9,* 1–12.

Bosetti, L. (2004). Determinants of school choice: Understanding how parents choose elementary schools in Alberta. *Journal of Education Policy, 19*(4), 387–405.

Bundorf, M. K., & Szrek, H. (2010). Choice set size and decision making: The case of Medicare Part D prescription drug plans. *Medical Decision Making: An International Journal of the Society for Medical Decision Making, 30*(5), 582–593.

Campos, S., Doxey, J., & Hammond, D. (2011). Nutrition labels on pre-packaged foods: A systematic review. *Public Health Nutrition, 14,* 1496–1506. doi:10.1017/S1368980010003290

Chakraborty, G., Ettenson, R., & Gaeth, G. (1994). How consumers choose health insurance: Analyzing employees' selection process in a multiplan environment identifies the trade-offs consumers make and the benefits that affect their decision making. *Journal of Health Care Marketing,* 14(1), 21–33.

Chen, Y., & Xie, J. (2005). Third-party product review and firm marketing strategy. *Marketing Science, 24*(2), 218–240.

Collins, A., & Snell, M. C. (2000). Parental preferences and choice of school. *Applied Economics, 32*(7), 803–813.

Corcoran, S. P., Jennings, J. L., Cohodes, S. R., & Sattin-Bajaj, C. (2018). *Leveling the playing field for high school choice: Results from a field experiment of informational interventions* (Working Paper No. 24471). Cambridge, MA: National Bureau of Economic Research.

Cronqvist, H., & Thaler, R. H. (2004). Design choices in privatized social-security systems: Learning from the Swedish experience. *American Economic Review*, 94(2), 424–428.

DeSanctis, G. (1984). Computer graphics as decision aids: Directions for research. *Decision Sciences*, *15*(4), 463–487.

Feldman-Stewart, D., Kocovski, N., McConnell, B. A, Brundage, M. D., & MacKillop, W. J. (2000). Perception of quantitative information for treatment decisions. *Medical Decision Making, 20,* 228–238.

Gelman, A. (2005). Analysis of variance: Why it is more important than ever. *Annals of Statistics, 33*(1), 1–53. doi:10.1214/009053604000001048

Gelman, A. (2006). Prior distributions for variance parameters in hierarchical models. *Bayesian Analysis, 1*(3), 515–533.

Gelman, A., & Hill, J. (2007). *Data analysis using regression and multilevel/hierarchical models vol. 1*. New York: Cambridge University Press.

Gelman, A., Hill, J., & Yajima, M. (2012). Why we (usually) don't have to worry about multiple comparisons. *Journal of Research on Educational Effectiveness, 5,* 189–211.

Glazerman, S. M. (2017). *Shopping for schools: Mapping choice architecture in the education marketplace* (Mathematica Working Paper Series, No. 59). Princeton, NJ: Mathematica Policy Research.

Glazerman, S. M., & Dotter, D. (2017). Market signals: Evidence on the determinants and consequences of school choice from a citywide lottery. *Educational Evaluation and Policy Analysis, 39,* 593–619. doi:10.3102/0162373717702964

GreatSchools. (2013). *How do parents research and choose schools?* Retrieved from https://www.greatschools.org/catalog/pdf/How_Do_Parents_Research_and_Choose_Schools.pdf

Harris, D. N., & Larsen, M. (2015). *What schools do families want (and why)? School demand and information before and after the New Orleans post-Katrina school reforms*. New Orleans, LA: Education Research Alliance for New Orleans.

Hastings, J., & Weinstein, J. M. (2008). Information, school choice, and academic achievement: Evidence from two experiments. *Quarterly Journal of Economics, 123*(4), 1373–1414.

Hastings, J. S., Kane, T. J., & Staiger, D. O. (2009). *Heterogeneous preferences and the efficacy of public school choice* (NBER Working Paper No. 2145). Cambridge, MA: National Bureau of Economic Research.

Hibbard, J. H., Slovic, P., Peters, E., & Finucane, M. (2002). Strategies for reporting health plan performance information to consumers: Evidence from controlled studies. *Health Services Research, 37*(2), 291–313.

Jacobsen, R., Snyder, J., & Saultz, A. (2014). Information or shaping public opinion? The influence of school accountability data format on public perceptions of school quality. *American Journal of Education, 121*(1), 1–27.

Jin, G. Z., & Leslie, P. (2005). The case in support of restaurant hygiene grade cards. *Choices, 20*(2), 97–102.

Jochim, A., DeArmond, M., Gross, B., & Lake, R. (2014). How parents experience public school choice. Center for Reinventing Public Education. Retrieved from http://www.crpe.org/publications/how-parents-experience-public-school-choice

Johnson, E. J., & Goldstein, D. (2003). Do defaults save lives? *Science, 302*(5649), 1338–1339.

Jones, G., & Richardson, M. (2007). An objective examination of consumer perception of nutrition information based on healthiness ratings and eye movements. *Public Health Nutrition, 10*(3), 238–244.

Jurajda, Š., & Münich, D. (2010). Admission to Selective Schools, Alphabetically. *Economics of Education Review, 29*(6), 1100–1109.

Kassler, D., Nichols-Barrer, I., & Finucane, M. (2018). *Beyond "treatment versus control": How Bayesian analysis makes factorial experiments feasible in education research* (Working Paper No. 61). Princeton, NJ: Mathematica Policy Research.

Kelly, J. P., III, & Scafidi, B. (2013). *More than scores: An analysis of why and how parents choose private schools*. Indianapolis, IN: The Friedman Foundation for Educational Choice.

Klute, M. (2012). *Evaluation of Denver's SchoolChoice process for the 2011-2012 school year*. Denver: University of Colorado, Denver.

Loranger, H., & Nielsen, J. (2006). *Prioritizing web usability*. Berkeley, CA: New Riders Press.

Miller, J. M., & Krosnick, J. A. (1998). The impact of candidate name order on election outcomes. *Public Opinion Quarterly, 62*(3), 291–330.

National Center for Education Statistics. (2018, March). *Public charter school enrollment*. Retrieved from https://nces.ed.gov/programs/coe/pdf/coe_cgb.pdf

Neild, R. C. (2005). Parent management of school choice in a large urban district. *Urban Education, 40*(3), 270–297.

Parducci, A. (1965). Category judgment: A range-frequency model. *Psychological Review*, *72*(6), 407–418.

Peters, E., Dieckmann, N. F., Dixon, A., Hibbard, J. H., & Mertz, C. K. (2007). Less is more in presenting quality information to consumers. *Medical Care Research and Review, 64*(2), 169–190.

Quaschning, S., Pandelaere, M., & Vermeir, I. (2014). When and why attribute sorting affects attribute weights in decision-making. *Journal of Business Research, 67*(7), 1530–1536.

Schneider, M., Teske, P., & Marschall, M. (2000). *Choosing schools: Consumer choice and the quality of American schools.* Princeton, NJ: Princeton University Press.

Schneider, M., Teske, P., Marschall, M., & Roch, C. (1998). Shopping for schools: In the land of the blind, the one-eyed parent may be enough. *American Journal of Political Science, 42*(3), 769–793.

Schwarz, N., & Bless, H. (1992). Constructing reality and its alternatives: An inclusion/exclusion model of assimilation and contrast effects in social judgment. In L. L. Martin & A. Tesser (Eds.), *The construction of social judgments* (pp. 217–245). Hillsdale, NJ: Erlbaum.

Stan Development Team. (2016). rstanarm: Bayesian applied regression modeling via Stan. R package version 2.13.1 [Software]. Retrieved from http://mc-stan.org/rstanarm/

Stein, M., Goldring, E., & Cravens, X. (2010). *Choosing Indianapolis charter schools: Espoused versus revealed academic preferences*. Nashville, TN: National Center on School Choice.

Stewart, N., Chater, N., & Brown, G. D. (2006). Decision by sampling. *Cognitive Psychology, 53*(1), 1–26.

Suk, K., Lee, J., & Lichtenstein, D. R. (2012). The influence of price presentation order on consumer choice. *Journal of Marketing Research, 49,* 708–717.

Teske, P., & Schneider, M. (2001). What research can tell policymakers about school choice. *Journal of Policy Analysis and Management, 20*(4), 609–631.

Thaler, R. H., & Sunstein, C. R. (2009). *Nudge: Improving decisions about health, wealth, and happiness.* New York: Penguin Group.

Tversky, A., & Kahneman, D. (1991). Loss aversion in riskless choice: A reference-dependent model. *The Quarterly Journal of Economics, 106*(4), 1039–1061.

Valant, J. (2014). *Better data, better decisions: Informing school choosers to improve education markets*. Washington, DC: American Enterprise Institute.

Vessey, I. (1999). Cognitive fit: A theory-based analysis of the graphs versus tables literature. *Decision Sciences, 22*(2), 219–240.

Waller, R., & Duncan, D. (1969). A Bayes rule for the symmetric multiple comparisons problem. *Journal of the American Statistical Association, 64*(328), 1484–1503.

Whitehurst, G. J. (2017) Education Choice and Competition Index, 2016. Retrieved from https://www.brookings.edu/wp-content/uploads/2017/03/ccf_20170329_ecci_full_report.pdf

After the Main Text, References, and any Appendix that may be appropriate – and before the last cover, you will find the *About the series* page. This page always contains the text as outlined in the next section. There is no running head or foot on this page.

ABOUT THE SERIES

Policymakers and researchers require timely, accurate, evidence-based research as soon as it's available. Further, statistical agencies need information about statistical techniques and survey practices that yield valid and reliable data. To meet these needs, Mathematica's working paper series offers access to our most current work.

For more information about this paper, contact Steven Glazerman, Senior Fellow, at sglazerman@mathematica-mpr.com.

Suggested citation: Glazerman, Steven, Ira Nichols-Barrer, Jon Valant, Jesse Chandler, Alyson Burnett. "Nudging Parents to Choose Better Schools: The Importance of School Choice Architecture." Working Paper 64. Washington, DC: Mathematica Policy Research, November, 2018.