The Journal of Economic Education

ISSN: 0022-0485 (Print) 2152-4068 (Online) Journal homepage: http://www.tandfonline.com/loi/vece20

Using online compound interest tools to improve financial literacy

Edward Hubbard, Percival Matthews & Anya Samek

To cite this article: Edward Hubbard, Percival Matthews & Anya Samek (2016) Using online compound interest tools to improve financial literacy, The Journal of Economic Education, 47:2, 106-120, DOI: 10.1080/00220485.2016.1146097

To link to this article: http://dx.doi.org/10.1080/00220485.2016.1146097

Published online: 13 Apr 2016.

THE JOURNAL OF ECONOMIC EDUCATION
2016, VOL. 47, NO. 2, 106–120
http://dx.doi.org/10.1080/00220485.2016.1146097

RESEARCH IN ECONOMIC EDUCATION

Using online compound interest tools to improve financial literacy

Edward Hubbard[a], Percival Matthews[a], and Anya Samek[b]

[a]Department of Educational Psychology, University of Wisconsin, Madison, WI, USA; [b]Center for Economic and Social Research, University of Southern California, Los Angeles, CA, USA

ABSTRACT
The widespread use of personal computing presents the opportunity to design educational materials that can be delivered online, potentially addressing low financial literacy. The authors developed and evaluated three different educational tools focusing on interest compounding. In the authors' laboratory experiment, individuals were randomized to one of three display tools: text, linear graph, or volumetric graph. They found that the text and volumetric tools were most effective at improving understanding of interest compounding, whereas individuals using the linear tool made little gains. The superiority of the text over the linear tool runs counter to the prediction of theories that suggest advantages of graphics over text. For researchers, the authors' findings highlight the importance of pedagogy evaluation. For practitioners, they provide research-validated tools for online dissemination.

KEYWORDS
Compound interest; education intervention; experiment; visualization

JEL CODES
A20; D14

Less than one-third of the U.S. population comprehends the fundamental concepts of interest compounding or understands how credit cards work (Lusardi and Mitchell 2009; Lusardi and Tufano 2009). Nevertheless, understanding compound interest plays an important role in financial planning. For instance, individuals with an understanding of interest compounding are more likely to plan for retirement (Lusardi and Mitchell 2011a). In addition, individuals with low debt literacy incur higher fees when borrowing, have higher debt loads, and have difficulty judging their debt position; yet, they often do not seek financial advisors (Lusardi and Mitchell 2009). Although financial education may hold the potential to solve this problem, there is a dearth of proven educational methods, in large part because careful evaluations of different pedagogies are scarce.

Lusardi and colleagues (2014) argued that pedagogy and delivery matter a great deal for promoting learning in the personal finance domain. For example, some programs found that printed educational material had little effect on retirement plan participation, perhaps due to the hands-off nature of the delivery method (Bernheim and Garrett 2003; Bayer, Bernheim, and Scholz 2009). In contrast, others have found significant effects on behavior for interventions that involve in-person retirement seminars (Clark and D'Ambrosio 2008; Bernheim and Garrett 2003) or in-person financial education courses in high schools (Lusardi and Mitchell 2009; Bernheim, Garrett, and Maki 2001). However, hands-on, in-person financial education programs are costly, and individuals might not seek them out. Lusardi and colleagues (2014) and Heinberg and colleagues (2014) used online survey experiments with participants of the American Life Panel and found that videos, relative to written narratives, had substantial effects on short-term financial literacy measures, including in the area of interest compounding.

CONTACT Anya Samek anyasamek@gmail.com Center for Economic and Social Research, University of Southern California, 635 Downey Way, Los Angeles, CA 90089-3332, USA.
Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/vece.

The widespread use of personal computing alongside broad access to Internet connectivity presents new opportunities to design interactive educational materials and to deliver them online. Online tools offer a few noteworthy advantages when compared with face-to-face education: they are relatively costless to scale up to many users, and participation in using such tools requires less time investment on the part of the consumer. They also can be integrated into online courses at universities. Finally, these tools can be integrated into financial services Web sites to act as decision support during the decision-making process, which is important because many financial decisions are now made online.

Many investment and amortization calculators are currently available online. Given their sheer numbers, we did not seek to do a thorough review of the types of calculators that were available. However, it is clear that available online tools differ widely in their representation of compound interest. Two common types of tools available for showing the effects of compound interest are simple calculators with no graphical display, and visual displays with linear graphics.[1] Despite the popularity of such aids and the wide range that is available, we know of no empirical studies that have tested the comparative efficacy of these different types of tools. We propose that it is important to carefully evaluate the relative merits of different pedagogical approaches to determine what kinds of online tools are most effective for promoting financial literacy. Laboratory experiments, in which participants are randomized to receive one of several different tools, are the ideal way to infer the causal impact of different pedagogical approaches on financial literacy.

In this study, we developed and evaluated tools designed to support the concept of interest compounding. Understanding compound interest is particularly crucial for financial literacy in light of the practical applications both to retirement savings and credit card debt (Lusardi and Mitchell 2011b), which are of great relevance to the general population. We developed three different educational decision-support tools for interest compounding and evaluated them in a laboratory experiment with 86 undergraduate students. In the experiment, participants were randomized to use one of three display tools: text, linear graph, or volumetric graph.

To measure treatment effects, we designed financial literacy surveys related to interest compounding that had unique properties, which improved upon existing measures in the field. First, we developed distractor items that allowed us to disentangle general incorrect responses from those based on erroneous simple interest reasoning, which we hypothesized would be a prevalent misconception. Second, we elicited participants' reasoning by including free-response explanation items. Whereas multiple-choice items rely on inferring participants' reasoning, the explanation items can explicitly measure reasoning. The experimental approach allowed us to assess the causal impact of each tool on learning outcomes, and provided us with rich, individual-level data on participants' reasoning about compound interest.

Several theories highlight the potential advantages of graphical representations over textual representations for promoting learning (e.g., Hegarty 2011; Larkin and Simon 1987; Lurie and Mason 2007; Tversky et al. 2000; Winn 1989). These theories might arguably have predicted superior results from the linear and volumetric graphs relative to the text display. However, in our study, we did not find categorical support for the graphical representations. Instead, we found that the text and volumetric versions of the tool were most effective at improving financial literacy in the posttest, whereas students using the linear version continued to demonstrate commonly held misconceptions in their answers.

In the next section, we provide a background discussion. In the following section, we describe the tools that we developed and outline the experimental design. The subsequent section provides a summary of our results, and the last section concludes.

Background

As noted by Lusardi and Mitchell (2009), the delivery method of financial education may have an impact on learning outcomes. Despite the promise of using interactive learning tools for teaching interest compounding, the utility of these tools has yet to be empirically investigated. In this study, we developed and evaluated three different educational tools to teach interest compounding (i.e., text, linear, and volumetric). All tools were interactive and provided feedback to the user in real time, but they varied in the way that data were displayed. In the text tool, data were displayed in table format. In the linear tool, we used a linear graph that plotted time on the X-axis and account balance (principal plus interest) on the Y-axis. Finally, in the volumetric tool, we used squares in 2-D space to represent account balance and color to indicate time.

The design of our tools and our theoretical perspective were informed by literature expanding beyond the financial education field. We chose to incorporate graphical representations (i.e., linear and volumetric tools) because psychological research has shown that graphical representations can play a facilitative role in problem solving and knowledge acquisition across a variety of domains (e.g., Winn 1989; Larkin and Simon 1987). Graphical representations shift information processing to the perceptual system, enhancing problem-solving capabilities and helping the user to identify trends and access information (e.g., Lurie and Mason 2007). Moreover, graphics (particularly those that express dynamic relations) can directly illustrate information that is only implicit in static textual displays (Larkin and Simon 1987). These features of graphical representations may be especially important in the domain of financial literacy, given the cognitive load required to examine textual representations of financial data.

Laboratory experiments have recently been used to investigate the impact of different graphical data representations on decision-making in economic games. For example, Cason and Samek (2014) investigated the impact of a linear graphical display as compared to a text display on decision-making in an asset market context. The authors found that the text display resulted in larger asset market bubbles than the graphical display; individuals in the laboratory seemed to be able to avoid falling prey to bubbles when information was presented in a graphical format. Samek and colleagues (2015) also found that different interactive displays, relative to their static counterparts, promoted improved decision-making in an information search task.

Our main hypothesis, based on related empirical work and theory, was that the graphical tools (i.e., linear and volumetric) would lead to greater learning of compound interest concepts than the text tool would. We hypothesized that graphical tools would prove more useful than the text tool for three reasons: first, we expected that the visuospatial representations of magnitude used by the graphical tools would communicate clearer semantic content to users than would the numerical symbols of the text tool. We suspected this because of well-documented and widespread problems with numeracy. Second, we suspected that linear and volumetric tools would reduce cognitive demand by recruiting perceptual abilities. Specifically, we predicted that graphics would allow processing of trends with a simple scan of the visual field, whereas the text tool would require keeping track of various point estimates and coordinating among them to extract information about trends.

The question of which type of graphical representation (linear or volumetric) would best promote knowledge of compound interest is more difficult to answer. The potential advantages of graphical representations must be qualified by the extent to which those representations align with the task at hand (Bassok, Chase, and Martin 1998). This is a particularly important consideration when predicting the relative efficacy of alternative representations that contain equivalent information for supporting decision-making processes (Larkin and Simon 1987). Research has shown that different graphical displays of the same information can lead to large differences in performance (e.g., Gattis and Holyoak 1996; Hegarty, Canham, and Fabrikant 2010; Tversky 2011; see Hegarty 2011 for a review). Hegarty (2011, 461) succinctly summed up the case thusly: "empirical studies have made it clear that one should not rely on intuitions alone to judge the effectiveness of visual displays."

In fact, in our experiment, decision makers can choose from multiple strategies to come to their answers, and the path they take might determine which tool is most effective. Moreover, a learner's prior knowledge of a given representational format might tilt the balance in its favor relative to others (Tversky 2011). For example, on the one hand, a decision maker can think in terms of graphical growth trends that are presumably best fit by linear graphs. On the other hand, decisions can be thought of in terms of comparisons of magnitudes at different time points, which might be best supported by our volumetric displays (if one focuses on area as an indicator of magnitude or amount). Thus, we did not have a clear-cut a priori hypothesis for the relative impact of linear versus volumetric tools. Note that similar considerations also may apply to determining the efficacy of text displays. Thus, our hypotheses about the relative efficacy of text displays have a modicum of uncertainty.

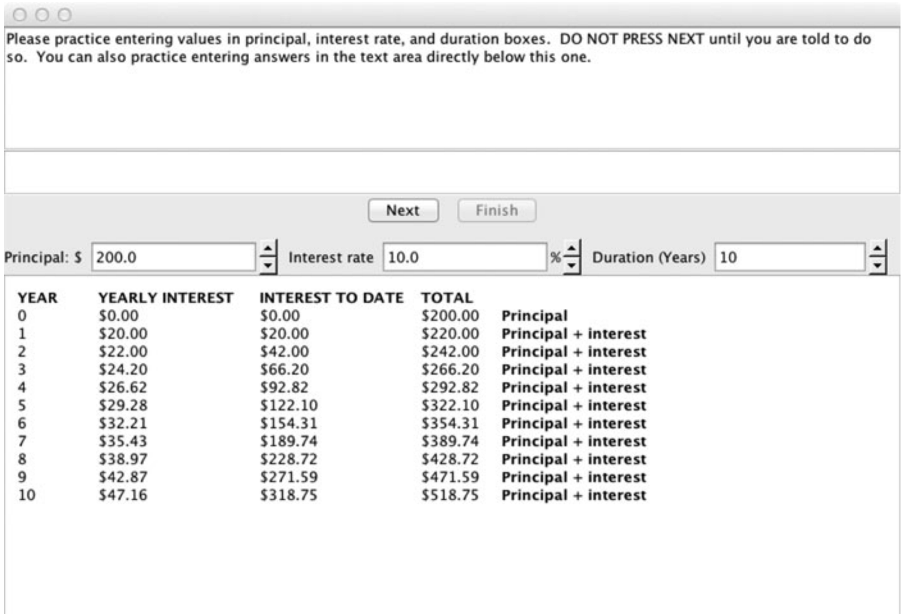

Figure 1. Text tool screenshot.

Educational tools and experimental evaluation

The educational tools

We developed three educational tools, a text tool, and two graphical tools (one linear and one volumetric). All tools were interactive and allowed users to investigate "what if" scenarios by manipulating the interest rate of the investment, the duration of an investment, and the initial principal invested. As the user changed the principal, interest rate, and duration (using either up/down arrows or by inputting a number), the tools provided information about the annual interest in each year, total interest to date, and the total balance at all points in the investment process. When introduced, each tool was accompanied with a set of detailed instructions that informed users on how to identify the principal and interest components.

The text tool, displayed in figure 1, provided the information as text in columns. The text tool required a greater amount of integration to process trends as compared to the graphical tools in that it required step-by-step scanning across years and a further synthesis step to conceive of overall trends. At the same time, it provided exact information at each time point.

The linear tool provided the same basic information as the text tool, only it used a line graph with years represented on the X-axis and balance in dollars on the Y-axis, and indicated the principal investment with a horizontal dotted line (figure 2). Thus, rather than performing calculations on number symbols across years to see trends (as was required with the text tool), the user could attend to the line representing total amount of interest and observe trends as inputs were adjusted. In this way, participants could get an intuitive feel for the trends associated with different interest rates and investment horizons by directly observing a linear or concave function and without paying explicit attention to the dollar values involved. However, determining interest accrued in a single year or the change from one year to the next requires coordination between the line graph and both axes. Thus, we hypothesized that analysis of year-to-year changes may be more labor intensive for users of the linear tool than it was for users of the text tool.

The volumetric tool also provided the same information about compounding interest, but represented this information using 2-dimensional boxes of different colors, which were added or removed with changes to principal, interest, and investment duration (figure 3). This tool used area as a perceptually accessible indicator of dollar amounts, and color allowed for quick identification of year-to-year

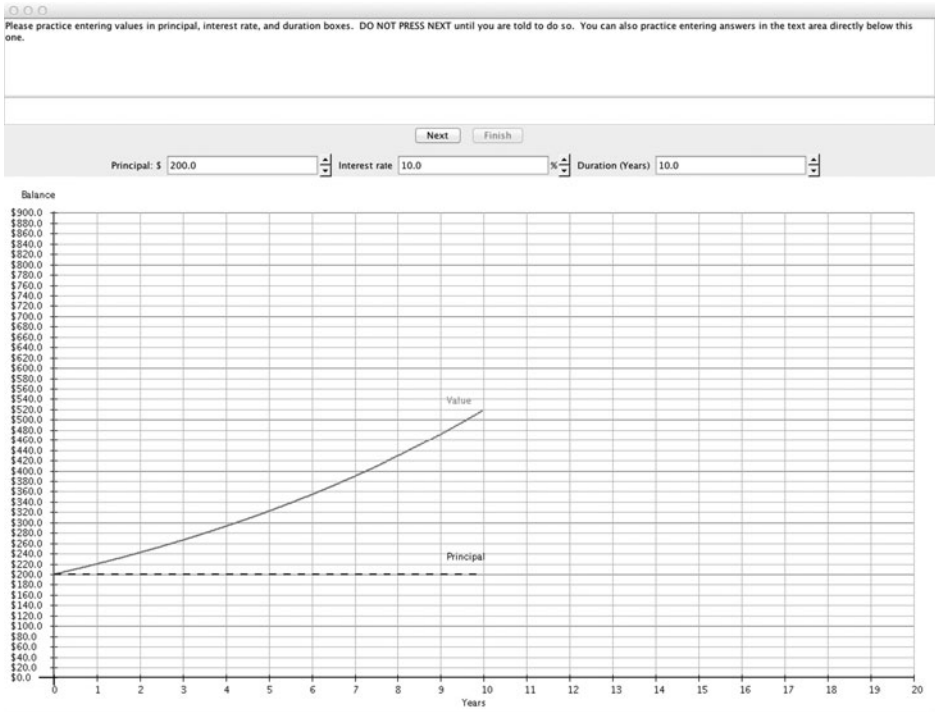

Figure 2. Linear tool screenshot.

changes. Like the linear tool, the volumetric tool also could show general patterns without requiring users to perform symbolic operations on actual dollar amounts. However, the volumetric tool lacked the linear tool's ability to show the nature of long-term trends (i.e., linear, concave upward).

Figure 3. Volumetric tool screenshot.

Experimental evaluation

Our experiment was conducted at the University of Wisconsin-Madison Behavioral Research Insights through Experiments (BRITE) Laboratory. The experiment was approved by the University of Wisconsin-Madison's Institutional Review Board. Participants were 86 undergraduate students (64 female; *mean age* = 20.1, *range* = 18–24) recruited from introductory educational psychology subject pools for course credit. Participation was voluntary, and students signed up online to participate in one of the offered sessions. Because there is no undergraduate degree in educational psychology at the university, majors varied broadly (nursing was the most common at 20.9 percent of the sample).

Participants were assigned to one of the three treatments: the text tool ($n = 24$), the linear tool ($n = 28$), or the volumetric tool ($n = 34$), and each participant used only one type of tool throughout the experiment. Assignment was blocked at the group level: everyone attending the lab for a given session was assigned to the same treatment at one of 20 available workstations. Thus, the imbalance in the number of observations across treatments was due to the number of participants who attended each session. Because students were not made aware of the treatment to which they would be assigned prior to signing up, or that multiple conditions existed, the decision to participate was not expected to be correlated with treatment. Also, all participants who came for the experiment decided to participate (i.e., no one opted out after reading the consent form or hearing the instructions).

Other than the differences in the type of tool used, the experimental procedure was identical across all three treatments. Each participant completed a single experimental session lasting approximately one hour. Students received credit in a course in educational psychology for participating in the experiment.

Each session employed a pretest-intervention-posttest design. Upon arrival, participants were assigned to a private computer station. The experiment proceeded in several parts. First, participants completed a pretest assessing their understanding of interest compounding without the aid of a tool. Second, we briefly explained to participants how to use and interpret their assigned intervention tool. Participants then completed an additional series of exploration questions involving compound interest, using their assigned educational tools to find the answers. Third, participants completed a posttest on interest compounding without the use of the educational tool. Finally, participants completed a short questionnaire collecting demographic information and inquiring about their financial education experiences. Students granted us access to their university math entrance examination scores (algebra, trigonometry, and basic skills subtests) as part of the consent process.

The pretests and posttests were intended to assess participants' unaided abilities to reason about compound interest before and after intervention with the learning tools. Each test included seven items requiring calculations involving compound interest problems, as well as two open-ended items prompting students to provide reasoning for some of their answers. Full versions of the pretest, posttest, and exploration items are available from the authors.

We designed the assessments with four goals in mind:

(1) We wanted to provide several items assessing compound interest to allow ample opportunity to observe variation in performance. This stands in contrast to the standard in prior literature of using a few questions (sometimes a single item) to measure understanding of compound interest (e.g., survey questions from the Health and Retirement Study).[2]

(2) We wanted to build upon established measures of compound interest knowledge. To this end, some of our compound interest investment questions mirrored the financial literacy questions used in related work (e.g., Lusardi and Mitchell 2009).

(3) We wanted to check for the prevalence of specific misconceptions, namely the use of simple interest logic when solving compound interest problems. Thus, we included naïve answers for multiple-choice items that corresponded to erroneous use of simple interest logic. For instance, for the first item shown below, application of simple interest to the original principal without accounting for compounding would lead to the selection of answer choice of 7 years instead of the correct answer of 5. For the second item shown below, failure to account for compounding would also lead to the conclusion that the yields were the same for both alternatives.

If you invest at a 15% interest rate compounded annually, about how long does it take for your investment to double? [INTEREST + PRINCIPAL]
a. About 1 year
b. About 5 years
c. About 7 years
d. About 15 years

Maribeth has a choice of investing $200 at a rate of 10% compounded annually for 10 years or $200 at a rate of 20% compounded annually for 5 years. Which investment will earn more in interest? [INTEREST]
a. 5 years at 20%
b. 10 years at 10%
c. They have the same yield
d. Not enough information was given

(4) We wanted to collect free response data on student reasoning to further investigate the cognitive processes that underpin the numerical accuracy scores. Thus, for two questions on both the pretest and posttest, we included prompts asking participants to "please explain the method you used to get your answer" and to "please explain why your method is correct."

The exploration questions were intended to provide participants with the opportunity to use tools to aid them in constructing understanding of interest compounding. There were 12 exploration questions, presenting a range of scenarios involving compound interest. Students were encouraged to use their assigned interactive learning tools to investigate each of the exploration questions. No feedback was given as to participants' accuracy until the end of the entire experiment. The exploration questions are available from the authors.

Coding free response items

For the free response items, we coded the rationale put forth by participants in terms of whether they used simple interest logic (SI) either in explaining their methods or in justifying why their chosen methods were correct. Coding was made solely on self-reported logic independently of whether participants actually selected the proper multiple-choice answer in the original question stem. Each pair of explanations (i.e., participants' descriptions of methods used and their justifications) served as a single item and was initially scored dichotomously as 0 or 1 (1 = SI).

Responses received an SI code only if participants explicitly indicated that they repeatedly applied simple interest on the original principal without considering compounding that accounted for accrued interest (see table 1 for scoring criteria and sample items). Responses that were ambiguous yet compatible with simple interest logic did not receive an SI code (SI = 0). After initial coding, scores for Q4 and Q7 were summed separately for pretest and posttest, yielding pretest logic and posttest logic scores for analysis. All responses were coded independently by two separate raters with 88 percent agreement.

Results

Overview of participants

As expected, participants were balanced across treatments based on key observables that we collected: gender, entrance exam subtests, and GPA (see table 2). Analyses of variance (ANOVAs), showing how each variable varied by condition, indicated that groups were statistically equivalent on each of the observable characteristics. Each of the entrance exam subtests was normed by the university to a mean of 500, so participants in our sample were average to above average in mathematical ability when compared to the university enrollment at large. Note that entrance exam scores were missing for seven participants because transfer students do not take the university system's entrance examination.

Table 1. Sample coding for free responses to questions 4a and 4b.

	If you invest at a 15% interest rate compounded annually, about how long does it take for your investment to double? [INTEREST + PRINCIPAL]	
	Explain the method you used	Explain why your method is correct
Code		
Simple interest = 1	I multiplied 15% by 7 to reach over 100% which is when it would double	100*. 15 goes to 115 then keep going to you get to a total of 200
Applied explicit reasoning amenable to simple interest calculation.	I said about 7 years because compounded annually means it is adding another 15% every year so to reach another full amount from you starting investment you would need to reach 100% interest rate and 15×7 is 105% which is about how long it takes to double.	My method is correct because 15×7 = 105 and this is closer to doubling (100%) than 15×6 which is only 90%
Simple interest = 0		
Compound interest Applied some reasoning expressing appreciation of the fact that interest is added to the principal, making future interest calculations larger.	If you earn 15% interest on your investment each year, after the first year you will have 15%, after the second year you will have a little over 30%, after the third year you will have a little over 45%, after the fourth year you will have over 60%, and after the 5th year you will have at least 75%. After including the interest you made on the interest accumulation you have over 100% interest earned.	My method is correct because you have to take into consideration that the amount is compounded each time and use the new amount each time while figuring out the interest.
	I calculated how long it would take for 15% to come to total 100%. You have to add 15% each time to the total and then take 15% of that. You can't simply keep adding the same number.	Simple analytical approach of calculating value after each year to determine the year in which the investment reaches double the initial value.
Prior knowledge of algorithm	rule of 70	I used the equation (p+r)ˆt
Undefined	I did not know the answer so I guessed	My method is correct because I used an example to guess and check my results.
	I took a random number, and placed it into the equation for interest and saw how long it would take for the interest to double	I did not calculate the exact answers, so I am not entirely sure that I am correct. However, it made the most sense.

Table 2. Balance of participants across treatments.

	Linear group	Text group	Volumetric group
Percentage female	78.6	70.8	73.5
MSBC score	604.07 (21.20)	630.42 (17.60)	615.00 (22.87)
Algebra score	559.96 (18.82)	579.58 (17.28)	565.36 (18.96)
Geometry/Trig score	522.96 (20.51)	578.75 (24.18)	556.43 (22.08)
GPA	3.13 (.18)	3.29 (.08)	3.36 (.07)

Note: Standard errors in parentheses.

Financial literacy at pretest

To check for equivalence of compound interest knowledge across treatments at pretest, we performed two tests. First, we compared overall pretest accuracy across the seven calculation items. Then, we compared groups' likelihood of choosing naïve answers that correspond to erroneous application of simple interest logic (which we refer to as "foils").

There were no statistically significant differences in financial knowledge by treatment at pretest as indicated by raw accuracy (see table 3). In table 4, we provide accuracy separately for each item. It is clear that participants' knowledge at pretest was poor: the only item that participants answered correctly with high frequency was question 5, and this item could correctly be answered using simple interest logic because it involved only one compounding period (i.e., it did not have a foil).

Two of the items (questions 1 and 2) were answered at chance levels. Finally, although the remaining four items were answered correctly at levels that exceeded chance, none of the four were answered correctly by more than 50 percent of participants.

We predicted that participants would tend to pick answers consistent with misapplication of simple interest logic at pretest, and this indeed was the case. We analyzed the frequency with which participants picked the incorrect answer corresponding to naïve application of simple interest for each of the six questions for which the analysis was possible (see tables 5 and 6). Notably, participants in each condition picked the naïve foil far more frequently than would be predicted by chance for five out of six possible questions. In terms of raw counts, participants picked the naïve answer more frequently than the correct answer for all six of the items. This is strong evidence that participants had low prior knowledge of interest compounding at pretest. Further, it demonstrates that specific misconceptions involving simple interest logic were quite prevalent. There was no difference between groups' tendencies to select the naïve foil at pretest.

Table 3. Pretest percentage correct by treatment.

	Linear ($n = 28$)	Text ($n = 24$)	Volumetric ($n = 34$)
Percentage correct	.39 (0.04)	.41 (0.04)	.45 (0.04)

Note: Standard errors in parentheses. Using *t*-tests, no significant difference between Linear v. Text ($p = .65$), Linear v. Volumetric ($p = .27$), nor Text v. Volumetric ($p = .53$).

Table 4. Pretest percentage correct by question.

	N	Percentage correct	Standard deviation
Question 1	86	0.28	0.45
Question 2	86	0.26	0.44
Question 3	86	0.36	0.48
Question 5	86	0.84	0.37
Question 6	86	0.41	0.49
Question 8	86	0.44	0.50
Question 9	86	0.34	0.48

Note: Questions 4 and 7 were free response and therefore did not have correct answers, so they were omitted from this table.

Table 5. Pretest selection of naïve foils by question.

	N	Percentage of foils	Standard deviation
Question 1	86	0.55	0.50
Question 2	86	0.31	0.47
Question 3	86	0.51	0.50
Question 6	86	0.43	0.50
Question 8	86	0.45	0.50
Question 9	86	0.62	0.49

Note: Questions 4 and 7 were free-response, and question 5 did not have a foil; therefore, these questions were omitted from this table.

Table 6. Pretest selection of naïve foils by treatment.

	Linear ($n = 28$)	Text ($n = 24$)	Volumetric ($n = 34$)
Percentage of foil selected (SD)	.49 (0.26)	.53 (0.26)	.43 (0.31)

Note: Using *t*-tests, no significance between Linear v. Text, ($p = .64$), Text v. Volumetric ($p = .22$), or Text v. Volumetric ($p = .40$).

Table 7. Pre- and posttest percentage correct by treatment.

	Pretest	Posttest	Differences in percentage correct
Linear	.38	.45	.07*
Text	.41	.60	.19***
Volumetric	.45	.64	.19***

Note: Pretest and posttest accuracy by treatment repeated in columns 1 and 2. Column 3 provides the difference and *p* value from paired *t*-tests within treatment.
$^{*}p < .10$; $^{***}p < .01$

Learning outcomes

To assess learning outcomes, accuracy for items of a given type (e.g., pretest items, posttest items, naïve foils, etc.) was summed and then divided by the total number of items. This created continuous measures of average accuracy to support the parametric analyses used in the remainder of this section.

We first calculated repeated measures *t*-tests comparing pretest and posttest accuracy in each treatment. Participants in volumetric and text treatments showed significant gains in improvement. Those in the linear treatment did not show a significant gain but did show a trend toward improvement (table 7).

To confirm that these relations held when controlling for other variables of interest, we regressed posttest accuracy against treatment condition, pretest accuracy, all three math entrance exam subtest scores, GPA, and gender (table 8). Treatment was dummy-coded for text and volumetric conditions with the linear condition serving as the baseline. Gender was coded with female as the baseline

Table 8. Results from OLS regression on posttest accuracy.

	Standardized coefficients	sr^2
Constant		
Text dummy	.263**	.049
Volumetric dummy	.343***	.083
Pretest accuracy	.274***	.057
MBSC	.406***	.063
ALG	− .159	.008
TRG	.066	.002
GPA	.056	.003
Gender (1 = male)	.072	.004
$R^2 = .429$		

$^{*}p < .10$; $^{***}p < .01$

Table 9. Pre- and posttest percentage foils by treatment.

	Pre	Post	Differences of percentage foils
Linear	0.49	0.49	– 0.01
Text	0.53	0.24	– 0.29***
Volumetric	0.43	0.21	– 0.22***

Note: Pretest and posttest foil choices by treatment repeated in columns 1–2. Column 3 provides the difference and p value from paired t-tests within treatment.
***$p < .01$

(male = 1). Finally, the three math entrance subtests were Algebra (ALG), Trigonometry (TRIG), and Basic Number Skills (MBSC). Thus, the ordinary least squares regression was

$$posttest_{simple} = b_1 Text + b_2 Volumetric + b_3 pretest_{simple} + b_4 MBSC + b_5 ALG + b_6 TRIG + b_7 GPA + b_8 GENDER.$$

To facilitate comparison of relative strength for different variables, our discussion for this and all subsequent regressions will refer to standardized coefficients.

We note that although performances on the three math subtests were highly correlated, each was designed to measure somewhat different abilities, any of which may have affected performance on the task. Accordingly, we included each of the three variables in the regressions, despite the risk of multicollinearity. We provide squared semipartial correlation coefficients (sr^2) in all regressions that follow to illustrate the independent contributions of each subconstruct.

Only the experimental treatment, pretest accuracy, and basic number skills were significant predictors of posttest accuracy (see table 8). Interestingly, basic number skills made large independent contributions over and above other math skills measures, accounting for 14.6 percent of all variance explained by the model. Still, participants who explored scenarios using the volumetric ($\beta = .34$, $p < .01$) and text tools ($\beta = .26$, $p = .02$) demonstrated superior learning as indicated by performance at posttest when compared to those who used the linear tool. In sum, performance differences based on treatment remained significant even after controlling for key variables of interest.

We analyzed changes in simple interest misconceptions to further investigate the learning gains observed in overall accuracy scores. Note that one posttest item was formulated in such a way that the naïve simple interest calculation was not included as an answer choice (Question 6). Thus, there were only five naïve foil items available at posttest as opposed to six at pretest. We first performed repeated measures t-tests comparing frequency of foil selection at pretest and at posttest in each condition. Consistent with overall learning scores, those in the text and volumetric treatments showed significant decreases in foil selection from pre- to posttest, whereas those in the linear treatment did not (table 9).

To test how all variables of interest affected selection of foils, we regressed proportion of posttest foil responses against treatment, pretest foil responses, math entrance exam subtest scores, GPA, and gender (see table 10). The final OLS equation was

$$posttest\_foil = b_1 Text + b_1 Volumetric + b_1 pretest\_foil + b_1 MBSC + b_1 ALG + b_1 TRIG + b_1 GPA + b_1 GENDER).$$

Only treatment, pretest foil selection, and basic number skills were significant predictors (table 10). Participants who explored scenarios using volumetric ($\beta = -.39$, $p < .01$) and text tools ($\beta = -.36$, $p < .01$) were significantly less likely to use select simplistic foils at posttest when compared to those who used linear tools.

Table 10. Results from OLS regression on posttest foil selection.

	Standardized coefficients	sr^2
Constant		
Text dummy	− .363***	.094
Volumetric dummy	− .390***	.109
Pretest foil selection	.350***	.092
MBSC	− .305**	.036
ALG	.069	.001
TRG	.017	.000
GPA	− .067	.004
Gender (1 = male)	.003	.000
$R^2 = .402$		

$^{**}p < .05$; $^{***}p < .01$

Analysis of free-response reasoning

The analyses above inferred participant use of simple interest logic from multiple-choice item responses. Analysis of free-response questions allowed us to explicitly examine participants' logic in order to corroborate our interpretations of the multiple-choice answer patterns. We regressed summed posttest simple interest logic scores against treatment condition, pretest simple interest logic scores, all three math entrance exam subtest scores, GPA, and gender (male = 1). As above, condition was dummy-coded for text and volumetric treatments, with the linear as the baseline condition. Thus, the ordinary least squares regression was

$$posttest_{simple} = b_1 Text + b_2 Volumetric + b_3 pretest_{simple} + b_4 MBSC + b_5 ALG + b_6 TRIG + b_7 GPA + b_8 GENDER.$$

Results are reported in table 11.

As reported in table 11, condition and pretest tendency to give simple interest logic were significant predictors of posttest tendency to report simple interest logic. Participants who explored scenarios using volumetric ($\beta = -.218$, $p = .05$) and text tools ($\beta = -.272$, $p = .02$) were equally less likely to offer explanations characterized by simple interest logic at posttest when compared to participants who used the linear tool. These patterns are consistent with our interpretations of response patterns for multiple-choice items above. However, they diverge from the above findings in that participants' basic number skills failed to significantly predict the tendency to use simple interest logic.

We conducted supplemental analyses to investigate this finding, in part because none of the three math-achievement math measures dominated the others in terms of sr^2, indicating that effects of multicollinearity might suppress an effect due to shared variance among the three subtests. We re-ran the regression analysis three times, only including one subtest in each analysis. When added independently, each of these measures was a significant predictor, whereas none were significant when added all together. This implies that some general math ability factor shared by each influenced performance. Although this

Table 11. Results from OLS regression on posttest simple interest logic.

	Standardized coefficients	sr^2
Constant		
Text dummy	− .272**	.053
Volumetric dummy	− .218*	.034
Pretest simple interest reasoning	.312***	.089
MBSC	− .178	.013
ALG	− .104	.003
TRG	− .096	.003
GPA	.091	.008
Gender (1 = male)	− .097	.008
$R^2 = .379$		

$^{*}p < .10$; $^{**}p < .05$; $^{***}p < .01$

contrasts with above findings in which basic number skills proved to be more important than other math skills, it does accord with the reasonable expectation that math skills should influence performance on compound interest tasks.

Discussion and conclusion

We developed and evaluated the impact of a series of computerized educational tools (text displays, linear graphs, and volumetric graphs) aimed at explaining the fundamental concept of interest compounding. We found that financial knowledge at pretest was low overall and marked by application of naïve simple interest calculations, consistent with findings from prior research (e.g., Hogarth and Hilgert 2002; Lusardi and Mitchell 2007, 2009). The text and volumetric tools proved superior to the linear tool in promoting learning as indexed by overall accuracy, and the linear tool failed to significantly improve overall accuracy. Moreover, the text and volumetric tools helped participants decrease their adherence to misconceptions based on naïve simple interest logic, whereas the linear tool did not.

These findings contribute to the field both by offering information about the comparative efficacy of different interactive support tools and by adding nuance to our understanding of the ways people process compound interest problems. One possible explanation for the positive impact of text tools, not predicted by our theory, might be found in analysis of decision strategies. If the strategy that is most helpful for learning was to find a point estimate rather than to recognize trends and patterns, this might explain why the text tool would have performed as well as or better than the graphical variants.

The poor performance of the linear tool relative to the others was particularly striking. We speculate (decidedly post hoc) about one factor that may have contributed to this lack of efficacy. Reading off dollar amounts on the linear tool required coordination between the X and Y axes to find the balance in an account at any given time point. By contrast, the text tool provided easy access to exact dollar amounts. The volumetric tool was organized into discrete squares that could be counted, so it did not require coordination across dimensions. Thus, it may have been that the linear tool, although informationally equivalent, was more cognitively taxing. One way to test this hypothesis in future work is to modify the linear tool such that pointing a cursor at different points provides a direct readout of the account balance and/or the elapsed time.

An additional innovation of the study was the use of a series of questions that allowed us to explore specific misconceptions ("foils") and to collect free-response data to learn about mechanisms. It is well-known that misconceptions pose barriers to learning in domains involving numeracy and that diagnosing misconceptions is important for promoting student improvement (e.g., Eryilmaz 2002; Resnick et al. 1989; McNeil and Alibali 2005). We showed that use of the volumetric and text tools reduced the likelihood of falling prey to misconception that simple interest logic should be used in compound interest calculation and that it helped remedy such misconceptions in a single session. Future work exploring whether the size of the effect increases with multiple learning sessions distributed across time will help determine the remedial potential of these tools.

Our findings have relevance both for research and for practice. For academics, we show (a) that the different ways to visualize information in online training tools have an impact on learning and the knowledge that is ultimately gained; (b) that specific misconceptions abound and that attending to these may prove to be particularly fruitful; and (c) that experimental evaluation of the impact of different pedagogies yields interesting and informative results for both theory and practice. For practitioners, we point to a new direction in developing educational tools that incorporate key concepts of interactivity and graphics to engage the user. For policymakers, who have been concerned about regulating content but have been less concerned about format, our findings on the different effects of each type of tool point to the importance of the role that format can play on understanding and decision-making.

Future work should investigate the broad potential of our tools for promoting financial literacy in general. One possible avenue for future research is to explore the use of these and similar online tools on a more diverse sample in order to investigate the effectiveness of the tools among the broader population.

In addition, future work should explore the use of interactive graphical tools for related financial literacy topics. In particular, it would be interesting to learn whether the efficacy of the tools extends to the realm of debt literacy.

Notes

1. Examples of tools with a nonvisual calculator include Investor.gov, the compound interest calculator for the U.S. Securities and Exchange Commission, and TheMint.org, a calculator provided by Northwestern Mutual to teach children and young adults about saving. Examples of tools with linear graphical representations include EconEdLink.org, an educational tool made possible by the Council for Economic Education.
2. http://hrsonline.isr.umich.edu/

Acknowledgements

The authors thank their research assistants at the University of Wisconsin-Madison for valuable assistance, especially Christa Gibbs, Tristan Abbott, Kathryn Carroll, and the Behavioral and Experimental Economics (BEE) research group.

Funding

The authors thank the UW-MadisonGraduate School for funding this research, which was carried out at the UW-Madison Behavioral Research Insights through Experiments (BRITE) Laboratory.

References

Bassok, M., V. M. Chase, and S. A. Martin. 1998. Adding apples and oranges: Alignment of semantic and formal knowledge. *Cognitive Psychology* 35 (2): 99–134.

Bayer, P. J., B. D. Bernheim, and J. K. Scholz. 2009. The effects of financial education in the workplace: Evidence from a survey of employers. *Economic Inquiry* 47 (4): 605–24.

Bernheim, B. D., and D. M. Garrett. 2003. The effects of financial education in the workplace: Evidence from a survey of households. *Journal of Public Economics* 87 (7–8): 1487–1519.

Bernheim, B. D., D. M. Garrett, and D. M. Maki. 2001. Education and saving: The long-term effects of high school financial curriculum mandates. *Journal of Public Economics* 80 (3): 435–65.

Cason, T., and A. Samek. 2014. Effect of visual representation and observational learning on asset market bubbles. Working Paper. West Lafayette, IN: Purdue University, Department of Economics, School of Management.

Clark, R. L., and M. D'Ambrosio. 2008. Adjusting retirement goals and savings behavior: The role of financial education. In *Improving the effectiveness of financial education and saving programs*, ed. A. Lusardi, 237–56. Chicago: Chicago University Press.

Eryilmaz, A. 2002. Effects of conceptual assignments and conceptual change discussions on students' misconceptions and achievement regarding force and motion. *Journal of Research in Science Teaching* 39 (10): 1001–15. doi:10.1002/tea.10054.

Gattis, M., and K. J. Holyoak. 1996. Mapping conceptual to spatial relations in visual reasoning. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 22 (1): 231.

Hegarty, M. 2011. The cognitive science of visual spatial displays: Implications for design. *Topics in Cognitive Science* 3 (3): 446–74.

Hegarty, M., M. S. Canham, and S. I. Fabrikant. 2010. Thinking about the weather: How display salience and knowledge affect performance in a graphic inference task. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 36 (1): 37.

Heinberg, A., A. A. Hung, A. Kapteyn, A. Lusardi, A. Savikhin Samek, and J. Yoong. 2014. Five steps to planning success. Experimental evidence from US households. NBER Working Paper No. 20203. Cambridge, MA: National Bureau of Economic Research.

Hogarth, J. M., and M. A. Hilgert. 2002. Financial knowledge, experience and learning preferences: Preliminary results from a new survey on financial literacy. *Consumer Interests Annual* 48 (1): 1–7.

Larkin, J. H., and H. A. Simon. 1987. Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science* 11 (1): 65–100.

Lurie, N. H., and C. H. Mason. 2007. Visual representation: Implications for decision making. *Journal of Marketing* 71 (1): 160–77.

Lusardi, A., and O. Mitchell. 2007. Baby boomer retirement security: The roles of planning, financial literacy, and housing wealth. *Journal of Monetary Economics* 54 (1): 205–24.

———. 2009. How ordinary consumers make complex economic decisions: Financial literacy and retirement readiness. NBER Working Paper No. 15350. Cambridge, MA: National Bureau of Economic Research.
———. 2011a. Financial literacy and planning: Implications for retirement wellbeing. NBER Working Paper No. w17078. Cambridge, MA: National Bureau of Economic Research.
———. 2011b. Financial literacy and retirement planning in the United States. *Journal of Pension Economics and Finance* 10 (4): 509–25.
Lusardi, A., A. Samek, A. Kapteyn, L. Glinert, A. Hung, and A. Heinberg. 2014. Visual tools and narratives: New ways to improve financial literacy. *Journal of Pension Economics and Finance* (2014):1–27.
Lusardi, A., and P. Tufano. 2009. Debt literacy, financial experiences, and over-indebtedness. NBER Working Paper No. 14808. Cambridge, MA: National Bureau of Economic Research.
McNeil, N. M., and M. W. Alibali. 2005. Why won't you change your mind? Knowledge of operational patterns hinders learning and performance on equations. *Child Development* 76 (4): 883–99. doi:10.1111/j.1467-8624.2005.00884.x.
Resnick, L. B., P. Nesher, F. Leonard, M. Magone, S. Omanson, and I. Peled. 1989. Conceptual bases of arithmetic errors: The case of decimal fractions. *Journal for Research in Mathematics Education* 20 (1): 8–27. doi:10.2307/749095.
Samek, A., I. Hur, S.-H. Kim, and J. S. Yi. 2015. An experimental study of information search and decision making with interactive technology. CESR-Schaeffer Working Paper. http://ssrn.com/abstract = 2347698. doi:10.2139/ssrn.23476982014.
Tversky, B. 2011. Visualizing thought. *Topics in Cognitive Science* 3 (3): 499–535.
Tversky, B., J. Zacks, P. Lee, and J. Heiser. 2000. Lines, blobs, crosses and arrows: Diagrammatic communication with schematic figures. In *Theory and application of diagrams: First International Conference, Diagrams 2000, Edinburgh, Scotland, UK, September 1–3, 2000 Proceedings*, ed. M. Anderson, P. Cheng, and V. Haarslev, 389–408. Springer-Verlag Berlin Heidelberg.
Winn, W. 1989. The design and use of instructional graphics. In *Advances in psychology, Vol. 58: Knowledge acquisition from text and pictures*, ed. H. Mandl and J. R. Levin, 124–44. Amsterdam: Elsevier B.V. doi:10.1016/S0166-4115(08)62151-2.