

Inflow Neglect: Forecasting Failures After Stocks Run Out

Megan E. Weber¹, Stephen A. Spiller¹, Hal E. Herschfield¹, and Suzanne B. Shu^{2, 3}

¹ Anderson School of Management, University of California, Los Angeles

² SC Johnson College of Business, Cornell University

³ National Bureau of Economic Research, Cambridge, Massachusetts, United States

People frequently encounter dynamic systems that involve inflows, outflows, and accumulated stocks—whether within their own households (e.g., financial accounts, stocks of food or supplies) or in larger institutional settings (e.g., manufacturing inventory, government benefit accounts). In this research, we introduce a novel stock-flow reasoning error, *inflow neglect*, and argue that this error can lead to important misperceptions regarding future outflows. To study this reasoning, we first focus on the United States' Social Security trust funds, whose impending depletion generates significant attention due to implications for American retirees. In Experiments 1–3, we show participants information about the trust funds over time that focus on the *stock* (i.e., balance) or *flows* (i.e., tax revenue and benefits payments), finding that those who see flows presentations are significantly less likely to expect benefits to cease completely after depletion (i.e., hold zero-outflow beliefs). In Experiments 4a and 4b, we show that prompting participants to reflect on the continuity of inflows (i.e., by reminding them that they expect payroll taxes to continue) significantly reduces inflow neglect and zero-outflow beliefs. Experiment 5 replicates these results in a separate domain, illustrating the generalizability of inflow neglect and underscoring the efficacy of presentations and targeted questions that emphasize the flows. This research contributes both theoretically and practically, advancing the literature on stock-flow reasoning and highlighting how communications about particular components of dynamic systems may contribute to—or be used to remedy—misconceptions that outflows will cease after depletion.

Public Significance Statement

People regularly encounter systems with inputs, outputs, and accumulation. Yet, when reasoning about the future of such systems, we find that people often ignore what will continue to flow into the system after the depletion of the stock, a bias we term *inflow neglect*. We highlight inflow neglect in a consequential public policy setting: Social Security benefits, a key source of financial stability for U.S. retirees. The agency's financial reserves are projected to become depleted within the next decade. When communications focus on the stock (i.e., the balance), large proportions of participants forecast that benefits will stop completely after depletion. However, graphs, written descriptions, or targeted questions that emphasize the system's flows and make the continuity of inflows more accessible combat inflow neglect and decrease the extent to which participants think benefits will stop completely. These results generalize to stock-flow reasoning beyond the Social Security setting.

Keywords: judgments, stocks and flows, retirement, public policy, communication

Supplemental materials: <https://doi.org/10.1037/xge0001891.supp>

Aysecan Boduroglu served as action editor.

Megan E. Weber  <https://orcid.org/0000-0002-2835-7939>

All study materials, preregistrations, code, and raw data are available on Research Box at <https://researchbox.org/1172>. The research reported herein was derived in whole or in part from research activities performed pursuant to the National Bureau of Economic Research (NBER; Grant RDR18000003 awarded to Stephen A. Spiller) from the U.S. Social Security Administration funded as part of the Retirement and Disability Research Consortium. The opinions and conclusions expressed are solely those of the authors and do not represent the opinions or policy of the Social Security Administration, any agency of the Federal Government, or NBER. Neither the U.S. Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or

usefulness of the contents of this report. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply endorsement, recommendation, or favoring by the U.S. Government or any agency thereof. The authors express gratitude to the University of California Los Angeles Anderson Morrison Center for Marketing and Data Analytics for their financial support, which made this research possible (Grant 23-014 awarded to the authors). As required by Retirement and Disability Research Consortium and NBER, earlier working article versions of this submission have been published as a part of conference proceedings for the 24th Annual Social Security Administration Research Consortium and as an NBER working article (NB22-15). The authors also thank attendees of the following seminars and conferences for their helpful comments in response to presentations of this work: University of California Los Angeles Behavioral Decision Making Lab meeting, 2022 Annual Meeting of the

continued

Understanding how accumulation works is a necessity of daily life. Inflows increase stocks (i.e., inventories, balances), and outflows draw them down. Households must track inventories of kitchen staples and toilet paper, monitoring use and ensuring inflows through shopping as outflows deplete the stock. Large organizations carefully track production, sales, and inventory to make sure that production rates are sufficient to meet the needs of customers as inventory waxes and wanes. Governments track increases and decreases in levels of water in reservoirs or levels of funds in financial accounts, paying close attention to ensure that constituent needs will be met. Yet stock-flow reasoning (i.e., translating between stocks and flows) in these types of situations is challenging for most people, such that errors happen regularly (e.g., Booth Sweeney & Sterman, 2000; Cronin et al., 2009; Sterman & Booth Sweeney, 2007). How do people forecast flows will change when the stock runs out? In this research, we find that a focus on the dwindling balance can lead to overly pessimistic forecasts of future outflows. We propose this results from a novel form of a stock-flow reasoning error we label *inflow neglect*: When a stock is nearing depletion, people tend to neglect the continuation of inflows and instead infer that outflows will stop completely when the stock runs out. In other words, inflow neglect leads to zero-outflow beliefs.

We use forecasts concerning the Social Security program's reserve accounts—often referred to as the “trust funds”—as a starting point to explore this broader issue. For American workers who anticipate receiving Social Security benefits in retirement, the solvency of the Social Security system is a pressing concern. In the absence of Congressional action, questions about the solvency of the system generally focus on the status of the Social Security trust funds. The trust funds are a textbook example of a dynamic system of accumulation, with inflows (i.e., payroll taxes paid into the system) and outflows (i.e., benefits paid out to beneficiaries) combining over time to determine the level of the stock (i.e., how much money is held in the trust funds). Recent projections are that the combined trust funds will be depleted by 2034 (The Board of Trustees, Federal Old-Age and Survivors Insurance and Federal Disability Insurance Trust Funds, 2025).¹ If and when the funds are depleted, Social Security is expected to be able to pay approximately 70%–80% of scheduled benefits. However, media coverage of the issue can paint an even more dire picture, with headlines that highlight how Social Security is “running out of money”—and many retirees are pessimistic about whether they will receive any benefits at all (Turner & Rajnes, 2021). Thus, the trust funds represent a practically important and theoretically interesting setting for our research on stock-flow reasoning in the face of impending depletion.

To begin to explore inflow neglect and its consequences for stock-flow reasoning, we present five experiments that vary the presentation format of information about the trust funds and examine how these different communications impact understanding. A sixth experiment establishes the generalizability of these findings to a

new, unrelated domain. We focus specifically on how stock versus flow information influences perceptions of what will happen to outflows as a result of depletion, building on work that highlights how small tweaks in the presentation of data and the formulation of questions can affect cognitive representations, understanding, and forecasts (e.g., Boz-Yılmaz & Boduroglu, 2024; Fischer & Gonzalez, 2016; Hoffrage et al., 2000; Hutchinson et al., 2010). We also identify factors that reduce inflow neglect and find that participants are less likely to jump to the conclusion that outflows will stop completely when they encounter graphs, written descriptions, or targeted questions that emphasize the flows (vs. the stock), suggesting that accessibility of system inflows may be a key driver of accurate understanding in stock-flow systems. Below, we provide a brief review of the literature on stock-flow reasoning problems, detail the application of this lens to the Social Security setting, and motivate how flows-focused communications and interventions may help attenuate inflow neglect.

Stock-Flow Reasoning

In the language of system dynamics, the cumulative amount of a resource, like the balance of money in a checking account or the amount of water in a reservoir, is a “stock.” The changes in the amount of a resource over some defined period, like deposits to and withdrawals from a checking account or water flowing into and out of a reservoir, are the “flows.” Given an initial value of the stock, there is a direct correspondence between the stock and the net flow: The stock is the integral of the net flow, and the net flow is the derivative of the stock. As a result, given either a time series of the stock or a time series of the flows (with a starting or ending value of the stock), the information content is calculably the same.

But calculus is challenging, and as a result, responses to the two representations often differ. First, formal mathematical transformations between stocks and flows are difficult and prone to error even among highly educated people (e.g., Booth Sweeney & Sterman, 2000; Brunstein et al., 2010; Cronin et al., 2009; Sterman & Booth Sweeney, 2007). This has sometimes been called the *stock-flow failure*. Because stock-flow transformations are so difficult, people are prone to rely on a faulty *correlation heuristic*: They focus too much on the specifics in front of them and wrongly expect the stock trend will tend to match the flow trend, rather than taking a broader view and considering the relationship between the elements (Cronin et al., 2009; Fansher et al., 2025; Fischer & Gonzalez, 2016). For example, a constant flow into a reservoir followed by a

¹ Separate projections exist for the Old-Age and Survivors Insurance (OASI, or retirement benefits) trust fund versus the smaller Disability Insurance (DI) trust fund, but the combined OASDI funds are the focus of most reports. Throughout the article and in our studies, we refer to the OASDI combined funds as “the trust funds.”

Retirement and Disability Research Consortium, 2022 and 2024 Society for Judgment and Decision Making conferences, and 2023 Boulder Consumer Financial Decision Making conference.

Megan E. Weber played a lead role in data curation, formal analysis, investigation, project administration, and writing—original draft and an equal role in conceptualization. Stephen A. Spiller played a lead role in supervision and writing—review and editing and an equal role in conceptualization and funding acquisition. Hal E. Herschfield played a supporting role in

conceptualization, supervision, and writing—review and editing and an equal role in funding acquisition. Suzanne B. Shu played a supporting role in conceptualization, supervision, and writing—review and editing and an equal role in funding acquisition.

Correspondence concerning this article should be addressed to Megan E. Weber, Anderson School of Management, University of California, Los Angeles, 110 Westwood Plaza, Suite B401, Los Angeles, CA 90095, United States. Email: megan.weber.phd@anderson.ucla.edu

constant flow out of a reservoir leads to a linear increase in the level followed by a linear decrease in the level. Instead, use of the correlation heuristic may lead people to infer that the level of the reservoir suddenly drops when the direction of flow shifts from in to out. Using such a heuristic can lead to both quantitative and qualitative mistakes, resulting in several serious errors like violating the conservation of mass in physical systems.

Beyond leading to these objective errors in calculation or translation, the presentation of stocks versus flows can also lead to qualitatively different evaluations and forecasts. Such *stock-flow inconsistencies* have been documented in several domains, including personal finances (Goldstein et al., 2016), evaluations of national employment (Spiller et al., 2020), and risk evaluations upon seeing COVID-19 test data (Reinholtz et al., 2021). As a concrete example, consider employment in the United States in 2009 (as tested in Spiller et al., 2020). During this time, the number of employed people was decreasing from one month to the next but at a slowing rate. The flow, or the change in the number of employed people month-to-month, was increasing (from a large negative number to a small negative number). But the stock, or the number of employed people, was still decreasing (due to the negative flow). As a result, a majority of respondents in the flow condition indicated that the economy was getting better while a majority of respondents in the stock condition indicated that the economy was getting worse. Presenting the same data in different ways thus led to qualitatively different evaluations. There were corresponding effects on forecasts of what respondents anticipated would happen next.

The literature on stock-flow reasoning failures and inconsistencies shows them to be remarkably robust and replicable for the translation from one format to the other, the interpretation of past data, and the projection of future trends. They apply across a wide variety of domains, including atmospheric accumulation of carbon dioxide (Serman, 2008; Serman & Booth Sweeney, 2007); water accumulating in a bathtub (Booth Sweeney & Serman, 2000; Cronin et al., 2009); weight (Brunstein et al., 2010); people in a store (Brunstein et al., 2010; Cronin et al., 2009); distance between cars (Cronin et al., 2009); product inventory in warehouses (Booth Sweeney & Serman, 2000; Spiller et al., 2020); national employment (Spiller et al., 2020); COVID-19 cases (Fansher et al., 2025; Harman et al., 2021; Padilla et al., 2022; Reinholtz et al., 2021; Villanova, 2022); blood glucose levels and other medical measurements (Brunstein et al., 2010); and perhaps most relevantly for the current investigation, corporate and personal cash flows (Booth Sweeney & Serman, 2000; Newell et al., 2016; Spiller et al., 2020). Such stock-flow failures and inconsistencies can be taken as a “stylized fact,” even among highly educated participants (e.g., medical students and Massachusetts Institute of Technology graduate students; Booth Sweeney & Serman, 2000; Brunstein et al., 2010; Cronin et al., 2009; Serman & Booth Sweeney, 2007). Moreover, they hold across multiple presentation formats, including scatterplots, line charts, bar charts, tables, and verbal descriptions (Cronin et al., 2009; Newell et al., 2016; Spiller et al., 2020).

An Exemplar of Problematic Inflow Neglect: Social Security Trust Funds

The financial future of the Social Security trust funds represents an important, interesting, and exemplary instance of problematic inflow neglect in stock-flow reasoning. The workings of the trust

funds map exactly onto stocks and flows: Over time, inflows (tax revenues) and outflows (benefits payments) combine to determine the stock or level of accumulation (trust funds balance). Each year, the Social Security Board of Trustees releases an Annual Report that documents actuarial forecasts for the future of the system. This report details forecasted tax revenue, benefits obligations, and trust funds balance in the short and long term. Some depictions focus on the balance of money in the trust funds (i.e., the stock, which may present a dire picture to consumers in that it shows depletion by 2034), and some show the changes in inflows and outflows over time (i.e., the flows, which may present a more optimistic view by emphasizing that there will still be taxes paid into the system that can be used to pay for benefits obligations).²

Media coverage of these reports that asks, “Will Social Security Run Out of Money?” (Paul, 2022) or highlights a “worst-case scenario” (Werschkul, 2021), and politicians who declare that Social Security will go “bankrupt” (Kiely, 2022) can make the actual impact on retirement benefits appear more severe than trust fund insolvency would actually imply. In reality, without Congressional action, approximately 70%–80% of benefits would still be payable after depletion (The Board of Trustees, Federal Old-Age and Survivors Insurance and Federal Disability Insurance Trust Funds, 2025).³ Instead, survey evidence suggests that many prospective retirees are pessimistic about future benefits. Turner and Rajnes (2021) found that many workers expect that they will not receive any retirement benefits at all (see also Luttmer & Samwick, 2018). This research suggests that current communications that focus on insolvency may be partly to blame and point to the need for targeted informational interventions to help improve accuracy.⁴ Quinby and Wettstein (2021) directly tested related ideas, looking at how variations in newspaper headlines for otherwise identical articles about projected insolvency can affect investing and benefit claiming intentions. Compared to a headline emphasizing a “long-term financing shortfall,” respondents who saw headlines highlighting that revenue will still cover three fourths of benefits were more likely to report accurate expectations regarding future monthly benefit size.

Whereas prior research has documented how various presentations of stocks or flows can affect translations between the two representations and other judgments about the state of the system, it has focused less on how different presentations might affect beliefs about inflows and forecasts about the continuation of future outflows (in the case of Social Security, the aspect of the system that corresponds to the core function of Social Security: to provide benefits for retirees). This in itself highlights a novel and important consequence of stock-flow failures and inconsistencies. Furthermore, we propose that pessimism regarding future benefits—and its link to

² For example, <https://web.archive.org/web/20250316155535/https://www.ssa.gov/OACT/TRSUM/index.html>

³ It is not clear exactly how changes to benefits would be applied (e.g., as an even 20%–30% cut across the board, varied cuts based on obligated amount).

⁴ Under an assumption that worker perceptions about system solvency distract from efforts to reform it, Holahan and Schug (2000) designed a training tool that explains the system to students of economics. Their teaching materials introduce flowcharts to track how money flows between payroll taxes from workers, Treasury, and retired beneficiaries. However, they do not report any results of empirical testing to show how successful their approach is at communicating the workings of the trust fund to student audiences, so it is unclear whether the tool can be used to reduce inaccurate inferences about the trust fund among a broader population.

alarmist coverage of the trust funds that focuses on impending depletion—may be explained, at least in part, by inflow neglect. Specifically, communications that make the presence of *inflows* more accessible in the context of a depleting stock may help people realize that *outflows* will continue in the future, while communications that emphasize the stock may lead people to focus on the draining balance and neglect to consider that flows can continue even when the balance is \$0. To use the analogy of a bathtub: When people are shown a bathtub that is rapidly draining and which will soon be empty, the most salient aspect of the system is that the level is falling, not that the tap is still on.

Attenuating Inflow Neglect

In this article, we explore how communications and interventions grounded in behavioral science research on stock-flow reasoning can help improve reasoning about dynamic systems facing impending depletion—like the Social Security trust funds. As previewed above, we argue that alternative presentations of the trust funds that deemphasize the balance and instead focus on the continuity of inflows and outflows may help to lead to more accurate judgments and inferences. As a further theoretical contribution, we propose a novel intervention that combats inflow neglect by directly targeting the accessibility of future flows. This intervention helps to overcome the assumption that stock depletion implies a complete end to outflows by simply prompting people to reason through their existing understanding of how the system works, without providing any new information. Notably, the literature on such stock-flow problems has found that reasoning errors arising from stock versus flow presentations are stubbornly persistent, even in the face of simpler presentations, clearer graphs, more familiar contexts, feedback about incorrect initial answers, and incentives to answer correctly (Brunstein et al., 2010; Cronin et al., 2009; Newell et al., 2016). Our intervention takes a different approach to improve reasoning in this setting: Rather than targeting the understanding or applications of the mathematical relationships between stocks and flows, we simply direct participants' attention to one aspect of the system (the inflows) that they already know about and, without providing additional information, ask them to think about what would happen to the outflows when the stock is depleted.

In proposing why this intervention might improve reasoning about future outflows, we build on work that suggests changing the accessibility of prior knowledge can affect the extent to which that information is used to make choices or draw conclusions (e.g., Feldman & Lynch, 1988; E. J. Johnson et al., 2007; Weber et al., 2007). Within the context of Social Security funding, workers' paychecks reliably include automatic deductions for the payroll taxes that fund the Social Security system, so most people are likely aware of their own contributions to the system's inflows. However, even if people understand the connection between these deductions and the benefits paid to retirees, they may not have actively considered whether they will continue paying these taxes in the future once the trust funds are depleted. Thus, the belief that benefits payouts will cease after depletion may result from the information that is accessible (e.g., trust fund depletion) without incorporating the information that is implied but not readily accessible (e.g., expected continuation of payroll taxes). In other words, neglecting the continuity of inflows leads to zero-benefits beliefs. But if people already understand that inflows (taxes) connect to outflows

(benefits), making the continuity of those taxes more accessible may affect their perceptions of what will happen to benefits after the trust funds run out of money.

Thus, the present research explores how different presentation formats and question formulations affect key assessments regarding outflows after a stock runs out. We primarily focus on the future of Social Security's outflows, and we also test whether these insights apply beyond this specific setting. We expect that stock presentations may lead to better calibrated forecasts of when the stock balance will reach zero, an important milestone for policymakers and managers, whereas flows presentations may lead to better calibrated expectations of what will happen to outflows, which are important for constituents and customers. To investigate whether this is driven by inflow neglect, we test an intervention that prompts participants to reflect on whether inflows will continue. We hypothesize that such reflection will further reduce misunderstanding about what happens to outflows by increasing the accessibility of the fact that inflows will not stop when the stock is depleted.

Overview of Experiments

We conducted six experiments to test these research questions.⁵ The first five studies are focused on the important context of the Social Security trust fund, and the sixth addresses the generalizability of our findings to a completely separate accumulation context. We describe the methods and results of each experiment in detail below.

Transparency and Openness

For each experiment, we recruited participants from Amazon's Mechanical Turk (AMT) and used CloudResearch to set additional eligibility requirements for each study. We restricted our samples to adults located in the United States with 100+ approved Human Intelligence Tasks, an approval rating of 95% or above, and who had not completed any prior studies run for this specific project. All study designs and analyses were preregistered on <https://aspredicted.org> prior to data collection,⁶ and any additional data exclusions are described where applicable.⁷ We report results for all preregistered key measures in the sections below. Full model results are reported in Supplemental Section 4. We used R Version 4.0.0 for analyses. All raw data, analysis code, and research materials are available in our Research Box,⁸ and main manipulations and measures are reproduced in Supplemental Section 6. The two pilot studies reported in Supplemental Sections 1 and 2 (respectively) and the six experiments reported in the main text represent all data we collected

⁵ Data collection for this project was declared exempt by the UCLA Institutional Review Board.

⁶ All preregistrations are available in our Research Box at <https://researchbox.org/1172>.

⁷ For each experiment, we include participants who completed the study once in our analyses. Due to the mechanics of Qualtrics, it was possible for subjects to enter the study multiple times (i.e., after failing an initial screener, participants could try again). As a result, our raw data files include a small number of instances of the same workers entering multiple times. In Experiments 1, 2, 3, 4a, and 5, none of these participants completed the full study more than once so they are all included in analyses. In Experiment 4b, one participant completed the study twice; only the first response from this participant is retained for analyses.

⁸ See this link for Research Box at <https://researchbox.org/1172>.

regarding the relationship between stock-flow reasoning and the impending depletion of the Social Security trust funds (and, in the case of Experiment 5, the impending depletion of a firm's inventory).

Experiment 1

In Experiment 1, we examine an initial instance of inflow neglect in stock-flow reasoning to test whether information presented in terms of stocks (rather than flows) might exacerbate zero-outflow beliefs. To do so, we used stock versus flow stimuli based on data and presentations of those data used in the 2022 Trustees Reports and related communications, focusing on measures of objective understanding as our key dependent variables.

Materials and Procedure

Drawing on a pilot study,⁹ we aimed to recruit 1,000 participants from AMT to take this survey. One thousand forty-four unique participants started the survey, and after preregistered exclusions, our sample for analysis consisted of 1,001 participants who completed the study ($M_{\text{age}} = 41.96$, $SD_{\text{age}} = 13.16$ ¹⁰; 51% male, 48% female, 0.6% nonbinary, 0.5% prefer not to say¹¹). First, participants read the following brief description about the OASDI trust funds:

The Social Security Administration uses accounts called "trust funds" to store income that's collected through Social Security taxes and then eventually uses that money to pay out benefits. The OASI Trust Fund pays retirement and survivors benefits and the DI Trust Fund pays disability benefits, but the two are often referred to together as the OASDI Trust Funds. At the end of the year in 2021, the OASDI trust funds held \$2.85 trillion in total.

Social Security uses the OASDI trust funds to make benefits payments and pay administrative expenses. The trust funds receive income through two sources: tax revenue collected from workers and interest that comes from the investment of the money in U.S. Government securities. In 2021, Social Security's total income from both of these sources was \$56 billion lower than its total costs (benefits payments plus administrative expenses). This was the first time in many years that total income was lower than total costs. Social Security predicts that in future years, total income will continue to be lower than total costs. Because of this continued projected deficit, the trust funds balance is projected to reach \$0 at some time in 2035.

On the same page as the above description, participants were randomly assigned to see an accompanying graph that showed either the balance (*stock* condition, $n = 508$) or total income and expenditures (*flows* condition, $n = 493$) of the trust funds for the period 1994 through 2034. Historical data were shown on both graphs with solid lines, while projected data were represented with a dotted line. Both graphs reported the respective metrics in trillions of dollars.¹² These graphs are shown in Figure 1.

Next, we asked participants four key objective understanding questions: (a) When total costs did/will begin to exceed income (with a dropdown where participants could select any year from 1994 to 2050 or never); (b) when the trust funds did/will become depleted (with a dropdown where participants could select any year from 1994 to 2050 or never); (c) what will happen to benefits if trust funds are depleted (a multiple choice question with options for no benefits, smaller benefits, the same amount of benefits, or larger benefits); and (d) after depletion, what the monthly retirement benefits amount would be for

someone currently expecting \$1,000 per month (an open text box; asked only of participants who selected either the smaller benefits or larger benefits option for the prior question). The order of the first two questions (when costs began to exceed income and when the trust funds would become depleted) was randomized across participants. We included order as a factor in statistical analyses reported below. Importantly, whereas both key dates were explicitly mentioned in the text of the description, what would happen to benefits was not discussed, meaning that answering these latter questions required participants to make inferences regarding the outflow that went beyond the data that were explicitly provided. After this section, participants wrote a media headline about the trust funds,¹³ responded to questions about their own eligibility for benefits, and provided demographic information. The final page of the survey provided links to further resources about the trust funds and benefits calculations. Key measures used for all studies are shown in Table 1 and are reproduced in full in Supplemental Section 6 (the full survey materials are available in our Research Box).

Results

As specified in our preregistration, we coded answers to the first three objective understanding questions according to accuracy (see Table 2 for the breakdown of answers to each question by condition for Experiments 1–3).¹⁴ To analyze differences in accuracy by condition, we conducted three separate logistic regressions using contrast-coded predictors (-1 , 1) for condition, order of the first two questions, and their interaction. There was no significant difference across conditions in accuracy for answers to the question about when costs began to exceed income ($b = -0.00$, $z = -0.01$, $p > .99$), though a larger proportion of participants in the stock condition (80%) correctly identified when the funds were projected to become depleted, compared to the flows condition (72%; $b = 0.21$, $z = 2.87$, $p = .004$). Those in the stock condition were more likely to answer that benefits would completely go away (i.e., hold "zero-benefit beliefs") as a result of depletion (64%), compared to the flows condition (56%; $b = 0.17$, $z = 2.62$, $p = .009$).

Following our preregistration, we analyzed the fourth question (about the benefits amount someone expecting \$1,000/month would

⁹ A brief summary of results from this pilot study ($N = 403$) is available in Supplemental Section 1.

¹⁰ In all studies, we asked participants for their year of birth (as Social Security rules regarding retirement are based on this). Age calculations are made as if participants were born on July 1.

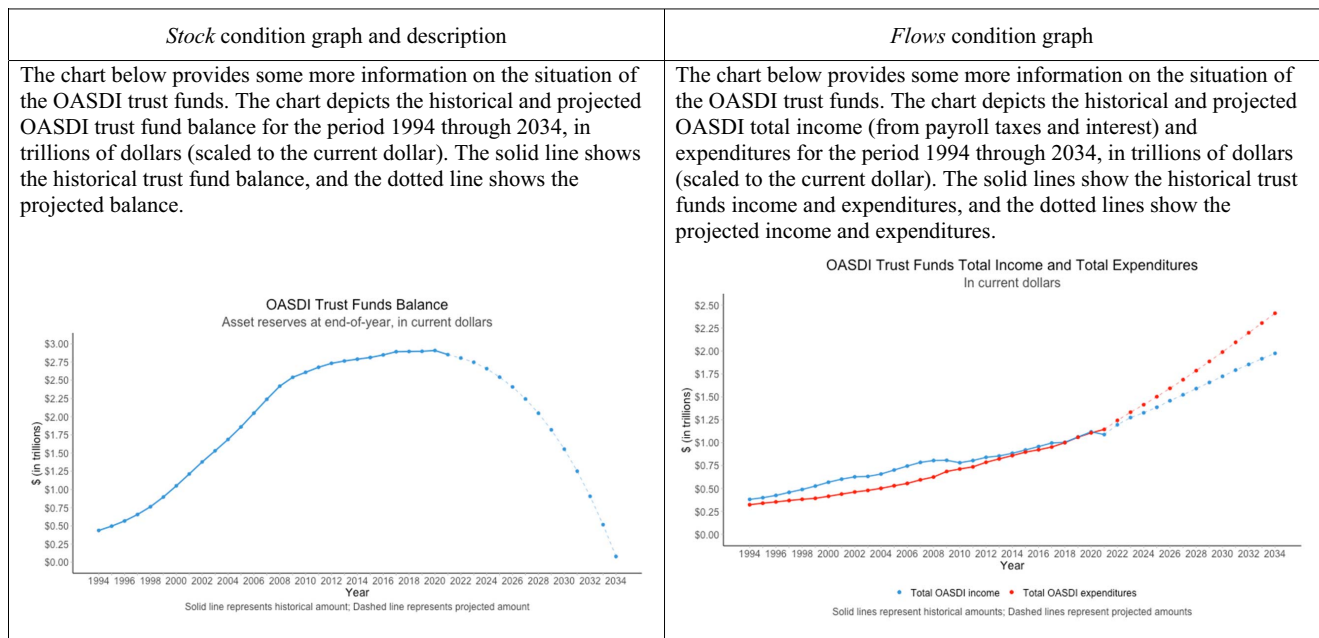
¹¹ For this and all following experiments, participants provided their year of birth, and age is calculated using that; the gender question asked "What is your gender" and gave respondents the option to select one option from the list.

¹² The description was based on the 2022 Trustees Report, and the data were taken from the 2022 Supplemental Single-Year Tables (specifically, Table VI.G8 available at <https://web.archive.org/web/20250222022501/https://www.ssa.gov/oact/TR/2022/lr6g8.html>).

¹³ We included this measure in Experiments 1, 2, and 4a to explore whether the different presentations led to systematically different "headlines" about the trust funds (i.e., calling out different aspects of the situation). A team of three research assistants coded these responses. There were no clearly interpretable patterns across conditions, so we do not discuss further. These data are available from the researchers upon request.

¹⁴ For all studies, missing answers due to a participant failing to answer a question are marked as missing but are included in the base for calculating the proportion of correct answers. These participants are excluded from regression analyses.

Figure 1
Experiment 1 Graph Stimuli (With Introductory Text)



Note. OASDI trust funds = combined Old Age and Survivors Insurance and Disability Insurance trust funds. See the online article for the color version of this figure.

get after depletion) as a continuous variable. This question was only asked of those who indicated in the prior question that benefits would be smaller or larger; we treated those who indicated that benefits would go away completely as giving an answer of \$0 to this question and those who indicated that benefits would stay the same as giving an answer of \$1,000. As preregistered, we also excluded participants who wrote in an answer above \$2,000. The average amount given by those who answered this question was \$641 ($SD = \238), but across the entire sample (i.e., using our imputed values), the mean was \$276 ($SD = \375). Using the same predictors specified above, we conducted a regression to analyze differences across condition. As expected, the average benefits amount was significantly lower for those in the stock condition ($M_{\text{stock}} = \$239$, $SD_{\text{stock}} = \$356$) compared to those in the flows condition, $M_{\text{flows}} = \$314$, $SD_{\text{flows}} = \$391$; $b = -37.00$, $t(996) = -3.13$, $p = .002$. To put these numbers into context, based on the projections in the 2022 Trustees Report, for every \$1,000 in scheduled benefits, trust funds income postdepletion would be sufficient to pay approximately \$750–\$800, meaning both groups still underestimated the implied amount on average.

We randomized the order of the first two questions that participants answered such that some participants saw the question about when depletion would happen first while others saw the question about when costs began to exceed income first. While this factor was included to counterbalance stimuli and was not of theoretical interest, results indicated that question order did have a significant impact on accuracy for the question about when costs began to exceed income ($b = 0.22$, $z = 3.38$, $p < .001$). Specifically, those who saw this question first were less likely to answer it correctly (51%) than those who saw the question about the depletion date first

(61%). We did not have a specific theoretical expectation for this result but return to it in the General Discussion section.

Discussion

Using comparable stock and flows graphs based on Social Security data, the results from this study show that different ways of showing trust funds information over time can significantly impact important facets of understanding. Specifically, we found that people shown a stock graph that displayed the trust fund balance over time (making particularly salient the decline to zero around 2035) were more accurate in their understanding of when depletion is projected to happen. This is in line with the existing literature on stock-flow reasoning—and notable given that this date was explicitly provided in the text description for both conditions. However, while the stock condition boosted accuracy on this question, those who saw the stock graph were *less* accurate in translating this information into an understanding of what impact trust fund depletion is projected to have on benefits, thus suggesting that inflow neglect may be heightened in the stock condition, compared to the flows condition.

Experiment 2

Experiment 2 was designed for three main purposes. First, we added a third “enhanced flows” condition, based on existing Social Security Administration (SSA) communications, that showed the same information on income and costs along with information on “payable benefits.” We included this condition to explore whether this additional information on payable benefits would further

Table 1*Key Measures for Experiments 1, 2, 3, 4a, and 4b*

Question description	Sample question text and answer options ^a	Correct answer for analyses
Costs exceed income date	According to your understanding, in what year did or will the Social Security OASDI trust funds' total costs begin to exceed total income? [Dropdown] 1994 ... 2050; Never	Experiments 1, 3, 4a, 4b: 2021 (± 1 year) Experiment 2: 2010 (± 1 year)
Depletion date	According to your understanding, in what year did or will the Social Security OASDI trust funds become depleted? In other words, in what year did or will the trust funds run out of money? [Dropdown] 1994 ... 2050; Never	Experiments 1, 2, 4a, 4b: 2035 (± 1 year) Experiment 3: 2034 (± 1 year)
What happens to benefits	Assuming the government does not take any action to increase the amount of income that Social Security collects based on tax revenues—in your view, what is most likely to happen to Social Security benefits if the trust funds are depleted? If you are not sure, please select the option that reflects your best guess. [Multiple choice (select one)] • Social Security will no longer be able to pay out benefits • Social Security benefits will still get paid out and will be a smaller amount • Social Security benefits will still get paid out and will be the same amount • Social Security benefits will still get paid out and will be a larger amount	Social Security benefits will still get paid out and will be a smaller amount
Benefits amount ^b	You just indicated that Social Security benefits will most likely still get paid out and will be a different amount if the trust funds are depleted. Again, assuming the government does not take any action to increase the amount of income that Social Security collects based on tax revenues—for someone whose benefits are currently projected to be \$1,000 per month, what monthly amount might they expect after depletion of the trust funds? \$[text box].00 per month	Analyzed as a mean (SSA projections imply the correct answer is around \$700–\$800.)

Note. SSA = Social Security Administration; OASDI trust funds = combined Old Age and Survivors Insurance and Disability Insurance trust funds.

^aThese are the exact questions used in most studies. Experiments 2 and 3 had small differences in the question text and correct answers based on what was shown in the graph and/or updated information from the Trustees. See [Supplemental Section 6](#) for full materials and questions from each study. ^bThis question was only asked of participants who responded “Social Security benefits will still get paid out and will be a smaller amount” or “Social Security benefits will still get paid out and will be a larger amount” to the prior what happens to benefits question. Per our preregistrations, answers to this question were imputed for participants who chose the other two answer options for the prior question, and we exclude participants who wrote an answer above \$2,000.

enhance understanding by making it clear when and how benefits would be impacted, thus representing an initial effort at further combating inflow neglect. Second, this study probes the robustness of the effect found in Experiment 1 by using graphs that show a

similar overall story but with different metrics and different axes. Third, we sought to investigate whether the findings from Experiment 1 would replicate with materials more closely based on those typically included in Trustees Reports, shifting from our

Table 2*Summary Table of Results for Main Preregistered Dependent Variables for Experiments 1–3*

Measure ^a	Experiment 1 (<i>N</i> = 1,001)		Experiment 2 (<i>N</i> = 1,503)			Experiment 3 (<i>N</i> = 1,501)		
	Stock	Flow	Stock	Plain flow	Enhanced flow	Stock	Flow	Net flow
Costs exceed income date (% correct) ^b	56%	56%	47%	46%	43%	54%	54%	53%
Depletion date (% correct) ^b	80%*	72%	74%	71%	73%	75%*	70%	70%
Zero-benefits beliefs (%) ^c	64%*	56%	61%*	54%	54%	60%*	54%	52%
Benefits amount, <i>M</i> (<i>SD</i>) ^d	\$239* (356)	\$314 (391)	\$264* (373)	\$318 (390)	\$321 (396)	\$270* (373)	\$337 (399)	\$323 (385)

Note. Asterisks (*) indicate a statistically significant difference ($\alpha = .05$) between the indicated condition and the preregistered reference condition for our main outcomes: Experiment 1 uses the flows condition as the reference category, Experiment 2 uses the plain flows condition as the reference category, and Experiment 3 uses the net flows condition as the reference category.

^aMissing answers due to a participant failing to answer a question are marked as missing but are included in the denominator for calculating the proportion of correct answers. These participants are excluded from regression analyses. ^bAs preregistered, answers are coded as correct if the participant entered a date within ± 1 year of the correct answer. ^cThe percentage of participants choosing this zero-benefits belief option is a key dependent variable across all studies. ^dFollowing our preregistration, responses above \$2,000 are excluded from calculations. The correct value is projected to be approximately \$700–\$800.

earlier stimuli that were generated to ensure formal equivalence across conditions.

Materials and Procedure

A total of 1,578 unique workers from AMT started the survey, and our sample for analysis included 1,503 participants who completed the study ($M_{\text{age}} = 40.90$, $SD_{\text{age}} = 12.91$; 47% male, 52% female, 0.6% nonbinary, 0.7% prefer not to say). The structure of the survey was very similar to that of Experiment 1: Participants saw a description and graph related to the trust funds, answered several objective understanding questions, and ended by answering questions about eligibility and demographics. The key changes in this study came in the data displays and content of the description. Participants were randomly assigned to see one of three graphs displaying information about the combined OASDI trust funds for the period 2000–2050: a *stock* graph that showed the trust fund ratios (balance as a percentage of projected costs for the ensuing year; $n = 500$), a *plain flows* graph that showed noninterest income and cost (scheduled benefits) as percentages of taxable payroll ($n = 502$), and an *enhanced flows* graph that built on the plain flows graph by including a line for expenditures (payable benefits; $n = 501$). We designed these graphs to mimic graphs from the Trustees Report as closely as possible, though we did constrain the x -axis to show the same date range across conditions to be consistent (see Figure 2 for our graphs and the SSA graphs we based them on). Since these stimuli were closely based on the graphs included in Trustees Reports, which use data on noninterest income (rather than total income, as we used in Experiment 1), we updated the description to include key dates and other information for this metric (specifically, the date for when costs began to exceed noninterest income is 2010; see Supplemental Section 6 or Research Box for full materials).

Results

As in Experiment 1, we preregistered that we would recode the objective understanding questions for accuracy and focus on proportions of answers reflecting beliefs that benefits will go away after depletion. Based on the three conditions in our design, we used the plain flows condition as the reference group and created two dummy-coded predictors to capture the effect of condition (stock vs. plain flows: 1 if stock condition, 0 otherwise; enhanced vs. plain flows: 1 if enhanced flows condition, 0 otherwise). As in Experiment 1, we randomized the order of the two questions about when costs began to exceed income and when depletion would happen and included a contrast-coded predictor ($-1, 1$) for order and all interactions in all analyses. Again, there were no significant differences across conditions for accuracy on the question about when costs began to exceed (noninterest) income (stock vs. plain flows: $b = 0.00$, $z = 0.00$, $p > .99$; enhanced vs. plain flows: $b = -0.16$, $z = -1.21$, $p = .22$).¹⁵ In contrast with our other experiments, there were no significant differences for the question about when depletion would happen (stock vs. plain flows: $b = 0.15$, $z = 1.08$, $p = .28$; enhanced vs. plain flows: $b = 0.12$, $z = 0.87$, $p = .38$), though the direction of the difference between the stock and two flows conditions is consistent with what we find in other studies.

We replicated the main results about understanding what would happen to benefits: Participants in the stock condition were more likely to choose the wrong answer about benefits going away

completely (i.e., hold zero-benefits beliefs) than those in the plain flows condition (61% vs. 54%; $b = 0.29$, $z = 2.22$, $p = .027$). Contrary to our expectations, however, the enhanced flows condition did not provide a further reduction in inaccuracy on this question, as the difference between the plain flows and enhanced flows condition was negligible and not significant (both 54%; $b = -0.01$, $z = -0.04$, $p = .96$).

We also analyzed the question about the amount of the expected benefits after depletion using the same method as Experiment 1. Among those who answered this question, the average amount entered was \$6230 ($SD = \253); across the whole sample, the average was \$301 ($SD = \387). Replicating Experiment 1, those in the stock condition ($M_{\text{stock}} = \$264$, $SD_{\text{stock}} = \$373$) thought benefits would be significantly lower than those in the plain flows condition ($M_{\text{plain flows}} = \318 , $SD_{\text{plain flows}} = \390 ; $b = -52.66$, $t(1495) = -2.15$, $p = .032$). However, the difference between the plain flows and enhanced flows condition was not significant ($M_{\text{enhanced flows}} = \321 , $SD_{\text{enhanced flows}} = \396 ; $b = 3.79$, $t(1495) = 0.15$, $p = .88$), providing further support for the conclusion that the mere inclusion of the “payable benefits” line did not necessarily improve understanding.

Discussion

In contrast to the prior study, neither question about dates showed any significant differences. More importantly, however, this study provides further evidence of the “stickiness” of the misconception that when the trust funds become depleted, benefits will fully disappear. Building on the previous study, we expected and found that the flows condition partially mitigates zero-benefits beliefs relative to the stock condition. Contrary to our expectations, however, the results from this study also suggest that including an explicit depiction of “payable benefits” in the enhanced flow condition does not provide any additional benefit for reducing inflow neglect to improve reasoning about future outflows.

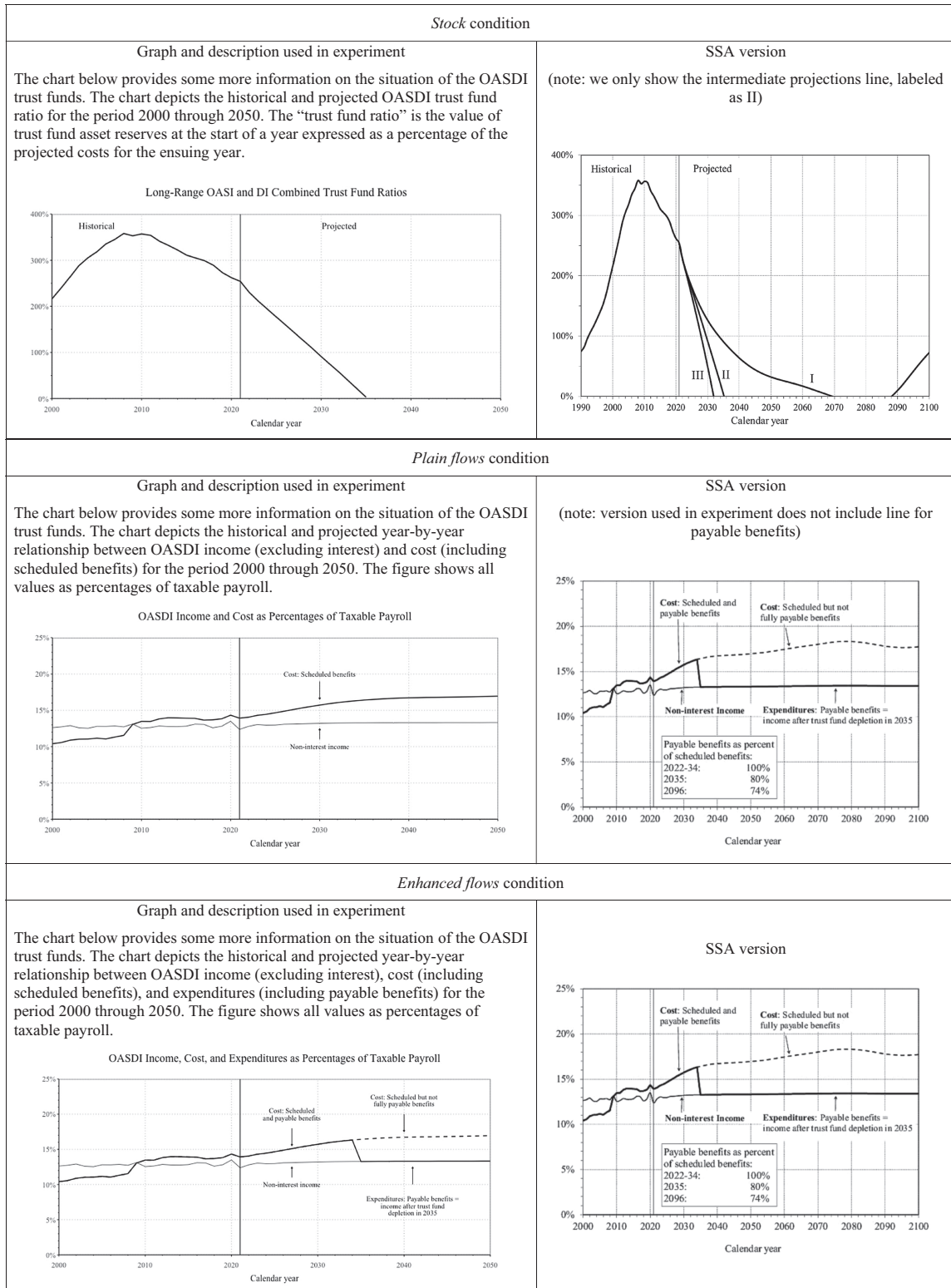
Experiment 3

We conducted Experiment 3 to explore the effect of an additional variation on presentation of information about the trust funds. Specifically, this experiment tested the same stock and flows presentations from Experiment 1 against a new condition that showed participants a graph of the net flows (i.e., the difference between income and expenditures over time).¹⁶ While Social Security does

¹⁵ There was, however, a significant interaction between the dummy code for stock versus plain flows and question order ($b = 0.37$, $z = 2.90$, $p = .004$). In particular, individuals were more accurate answering the question about when income began to exceed costs when the depletion question was first in the stock condition but more accurate when the income question was first in the plain flows condition. This interaction does not appear consistently in the other studies, and we do not have a specific theoretical explanation for it.

¹⁶ We also made two small text updates. First, we added one additional sentence to the trust funds description shown to all participants. This sentence provided participants with information about the current level of inflows and outflows as of the most recent Trustees Report, ensuring formal equivalence across conditions. The second change was to the key outcome question about what happens to benefits after depletion. For this question, we added a sentence to ensure that participants knew they should be answering about retirement benefits in general (rather than about their own retirement benefits). The full materials are available in Supplemental Section 6. As is evident in the results reported below, participants in this study provide answers that are very similar to what we observe in other studies.

Figure 2
Stimuli for Experiment 2, With Comparison to SSA Graphs



Note. The source for SSA graphs is found at https://web.archive.org/web/20250321170512/https://www.ssa.gov/oact/TR/2022/IL_D_project.html#105057. OASDI trust funds = combined Old Age and Survivors Insurance and Disability Insurance trust funds; SSA = Social Security Administration.

not typically report trust funds data in this way, inclusion of this condition in this experiment was useful for two main reasons. First, the pattern of net flows charted over time is visually similar to that of the stock graph in that it shows the net flows declining over time (see Figure 3). Past work on stock-flow reasoning errors has pointed to the correlation heuristic as an explanation for errors in translations between stocks and flows. According to a correlation heuristic explanation, people erroneously think the trend of the stock should match the trend of the net flows. If our results regarding zero-benefits beliefs were due to the correlation heuristic, we might expect a net flows presentation to result in a pattern similar to what we see in the stock condition. This is because both graphs show generally negative trends for the future (see Figure 3), whereas the flows condition shows costs and income increasing over time. On the other hand, if the net flows condition results look more like the flows condition, this would suggest that there is something unique about getting participants to think about the flows of the system and how they work together to determine benefits. Second, to address a potential alternative explanation described in more detail below, this experiment included one additional question about participants' baseline understanding of how the Social Security system works.

Materials and Procedure

A total of 1,620 unique workers from AMT started the survey, of whom 1,501 completed the study ($M_{\text{age}} = 41.90$, $SD_{\text{age}} = 12.39$; 46% male, 53% female, 0.7% nonbinary, 0.7% prefer not to say). The survey followed the same structure as prior experiments. First, participants read a description of the trust funds (updated as described below) and were randomly assigned to see a stock graph, a flows graph, or a net flows graph. The stock and flows graphs were the same as in Experiment 1 (with updated numbers based on the most recent Trustees Report¹⁷), and the net flows graph is shown in Figure 3 ($n_{\text{stock}} = 509$, $n_{\text{flows}} = 496$, $n_{\text{net flows}} = 496$). Next, participants answered the same four dependent variable questions and then responded to a new question about their mental model of the Social Security system. The survey concluded with questions about participants' eligibility for Social Security and demographic characteristics.

The new question about mental models of the Social Security system allowed us to test an important alternative explanation. Given the complexity of the Social Security system, it is possible that some participants have a fundamental misunderstanding of how the system works. Specifically, some people may not realize that current tax revenue is used to pay *current* beneficiaries (i.e., a "pay-as-you-go" kind of system), instead thinking that the system is more like a government-held individual savings account (such that an individual's payments into the system are literally deposited into an account and paid back to them once they retire). If at least some participants have this misconception, it is possible that manipulations focused on the flows simply correct this misunderstanding. In doing so, relatively more participants faced with flows may understand that future benefits are paid by future taxes, leading to a correct inference that as future taxes will continue, so will future benefits, whereas relatively more participants faced with the stock may (mis)understand that future benefits are paid by previous taxes, leading to an erroneous inference that once the trust funds are depleted, the funds to pay benefits have been exhausted. In other words, this alternative explanation hinges on stock versus flows presentations differentially helping participants understand the fact

that current inflows are related to current outflows. This is in contrast to the focal explanation that presenting flows simply makes inflows more accessible, allowing participants to use *what they already know* about the link between taxes and benefits to reason about the continuity of those flows after depletion. To address this possibility, we included a question after our main dependent variables that asked participants to report their understanding of Social Security as a pay-as-you-go system versus an individual savings account model.¹⁸ We examined how answers to this question were associated with perceptions about what happens to benefits.

Results

Like prior studies, across all conditions, about half of participants (53%) correctly identified that costs began to exceed total income in 2021, almost three fourths (72%) correctly identified the depletion date, and approximately one third (36%) correctly answered that benefits would still be paid out in smaller amounts after depletion. For the new question about understanding of how the Social Security system works, 77% of participants correctly identified that the system is pay-as-you-go.

Following our preregistration, we used dummy-coded predictors for condition (with the net flows condition as the reference category) to test how the net flows condition compared to the stock and flows condition on our key outcomes. We also conducted follow-up analyses using the flows condition as the reference category. As in prior studies, we randomized the order of the first two objective knowledge questions about key dates, so all models included a contrast-coded predictor for order (−1, 1) that controls for this factor and all interactions between factors. Supplemental Section 4 shows the full model results for all analyses.

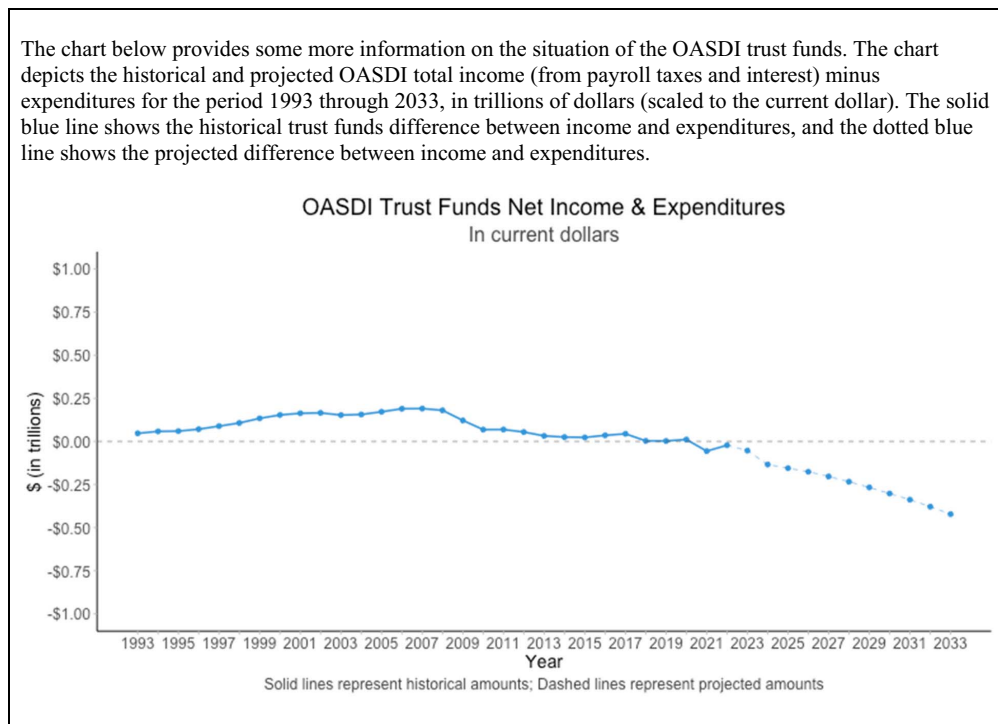
Replicating Experiments 1 and 2, there were no significant differences across conditions for the question about when costs began to exceed income (stock vs. net flows: $b = 0.05$, $z = 0.40$, $p = .69$; flows vs. net flows: $b = 0.05$, $z = 0.37$, $p = .71$), though we do find that those who saw this question *after* answering the depletion question were more likely to answer accurately ($b = 0.26$, $z = 2.84$, $p = .005$). In line with Experiment 1, those in the stock condition were more likely to correctly identify the projected depletion date (75%) compared to those in the net flows condition (70%; $b = 0.28$, $z = 1.99$, $p = .047$) and those in the flows condition (70%; $b = 0.30$, $z = 2.09$, $p = .037$).

For our main outcome question about what happens to benefits after depletion, fewer participants in the net flows condition chose the answer about benefits going away completely (53%), compared to the stock condition (stock: 60%; $b = 0.29$, $z = 2.26$, $p = .024$). The difference between the net flows and flows condition was not significant (flows: 54%; $b = 0.06$, $z = 0.44$, $p = .66$), and the difference between the stock and flows condition was marginally significant and directionally consistent with other studies ($b = 0.23$, $z = 1.81$, $p = .07$). Table 2 provides the full breakdown of answers to this question by condition.

¹⁷ This study was conducted after the release of the 2023 Trustees Report (available at <https://www.ssa.gov/OACT/TR/2023/tr2023.pdf>). The description and graphs were all updated to reflect the most recent data points and projections. The key difference is that the depletion date projected in this report is 2034.

¹⁸ Prior to running this experiment, we conducted a small pilot study to test this question. See Supplemental Section 2 for a brief summary of this pilot and its results.

Figure 3
Net Flows and Graph Used in Experiment 3



Note. OASDI trust funds = combined Old Age and Survivors Insurance and Disability Insurance trust funds. See the online article for the color version of this figure.

The pattern of results was similar for the question about the amount of benefits that would still be payable. Among those who answered this question, the average amount written in was \$634 ($SD = \244); using our preregistered approach to impute values based on answers to the prior question, the overall average was \$310 ($SD = \387). Looking at this by condition, the average amount in the net flows condition ($M_{\text{net flows}} = \$323$, $SD_{\text{net flows}} = \385) was not significantly different from the flows condition ($M_{\text{flows}} = \$337$, $SD_{\text{flows}} = \$399$; $b = 14.23$, $t(1493) = 0.58$, $p = .56$). The average benefits amount in the stock condition ($M_{\text{stock}} = \$270$, $SD_{\text{stock}} = \$373$) was significantly lower than in both the net flows and flows conditions

(versus net flows: $b = -52.47$, $t(1493) = -2.15$, $p = .032$; versus flows: $b = -66.71$, $t(1493) = -2.74$, $p = .006$).

We now turn to the question about how participants thought Social Security works, summarized in Tables 3 and 4. Limiting our analysis to only participants with the correct mental model (the majority of our sample), we still see that those in the stock condition were relatively more likely to say that benefits will go away completely (stock: 62%; flows: 54%; net flows 56%; stock vs. net flows: $b = 0.23$, $z = 1.55$, $p = .12$; stock vs. flows: $b = 0.32$, $z = 2.17$, $p = .03$). Thus, it does not appear that the flows condition reduced misperceptions regarding benefits by differentially correcting a misunderstanding that Social Security works like an individual's savings account. Instead, we contend the flows conditions helped participants better *use* their knowledge about how the system works to form predictions.¹⁹

Table 3
Mental Models of Social Security

Mental model	Overall proportion	By presentation condition		
		Stock	Flow	Net flow
Correct mental model	77%	74%	79%	78%
Incorrect mental model	23%	26%	21%	22%

Note. Participants in the stock condition were directionally less likely to choose the correct answer relative to the net flows condition ($b = -0.20$, $z = -1.37$, $p = .17$) and marginally less likely relative to the flows condition ($b = -0.28$, $z = -1.84$, $p = .066$). However, as described in the text, analyses of answers of our key outcome variable based on participants' mental model indicated that the effect of the flows condition does not stem from a differential correction of participants' mental models.

¹⁹ We also looked at whether those with the incorrect mental model (i.e., who think that Social Security works like an individual savings account) were more likely to think that benefits would go away completely after depletion. In fact, we find the opposite: Across our whole sample, those with the incorrect mental model are significantly less likely to have this misperception (49% vs. 57%; $b = 0.34$, $z = 2.75$, $p = .006$). Inspecting the full set of results, it appears this is accompanied by an across-the-board increase in the proportion of responses indicating benefits will be paid at the same level. While this is a post hoc interpretation, this suggests that the misconception that Social Security operates as an individual savings account may be associated with the inference "and therefore my funds are protected."

Table 4
Mental Models of Social Security and Benefits Beliefs

What happen to benefit question (% choosing each option) ^a	Correct mental model			Incorrect mental model		
	Stock	Flow	Net flow	Stock	Flow	Net flow
Benefits go away completely ^b	62%*	54%	56%	53%*	53%*	39%
Paid, smaller amount ^c	33%	37%	38%	35%	29%	49%
Paid, the same amount	4%	8%	6%	9%	13%	14%
Paid, larger amount	1%	1%	0%	2%	1%	1%

Note. Asterisks (*) indicate a statistically significant difference ($\alpha = .05$) between the stock condition and the flows condition for those with the correct mental model and between the net flows and the other two conditions for those with the incorrect mental model.

^a Missing answers due to a participant failing to answer a question are marked as missing but are included in the denominator for calculating the proportion of correct answers. These participants are excluded from regression analyses. ^b Whether or not participants choose this zero-benefits-beliefs answer option is a key dependent variable across all studies. ^c This is the correct answer.

Discussion

Experiment 3 contributes three takeaways. First, we confirm that our results do not change when those in the stock condition have information about current levels of income and expenditures (which, theoretically, should help them reason through why the stock is being depleted and what that means for future benefits) or when we make it more explicit that our main question about what will happen to benefits is asking about benefits in general. Second, we rule out an alternative explanation—namely, that the flows condition simply corrects a misunderstanding regarding their model of how Social Security works rather than helping people draw the correct implications from the correct (within reason) model. Finally, the fact that the net flows condition generates a pattern of results that closely mirrors that of the flows condition both rules out the correlation heuristic as an explanation and suggests that there is something unique about the accessibility of flows (vs. the overall downward trend apparent in both the stock and the net flows condition). The next two studies take this insight a step further by testing an intervention that attenuates inflow neglect by explicitly prompting participants to consider expectations about future inflows.

Experiments 4a and 4b

While the flows and net flows manipulations used thus far significantly decreased the extent to which participants thought benefits would go away completely, levels of misunderstanding remained consistently high—over half of participants in these conditions still exhibited zero-benefits beliefs. Experiments 4a and 4b were designed to test a stronger manipulation that encouraged participants to reason through what would happen to the inflows and outflows when the trust funds become depleted. This manipulation was inspired by Benjamin Franklin's famous utterance that "nothing can be said to be certain, except death and taxes" (Smyth, 1907, p. 69). Specifically, the intervention asked participants to consider whether Social Security will continue collecting taxes (in line with Benjamin Franklin, a large majority say "yes") and if so, what that money will be used for.²⁰ Experiment 4a provided an initial test of this novel intervention, and Experiment 4b served as a higher power replication of both the intervention effect from Experiment 4a and the stock-flow effect from Experiments 1 to 3. As the procedure and analyses are almost identical for the two studies, we describe them together below.

Materials and Procedure

Our final samples included 1,001 participants (out of 1,044 who started the study) from AMT for Experiment 4a ($M_{\text{age}} = 41.17$, $SD_{\text{age}} = 12.59$; 45% male, 52% female, 2% nonbinary, 0.3% prefer not to say) and 2,001 participants (out of 2,126 who started the study) from AMT for Experiment 4b ($M_{\text{age}} = 41.00$, $SD_{\text{age}} = 12.78$; 42% male, 55% female, 2% nonbinary, 0.5% prefer not to say). The beginning of the survey was identical to previous studies, where participants read a description of the trust funds accompanied by a stock or flows diagram (the same description and graphs as in Experiment 1; 4a: $n_{\text{stock}} = 499$ and $n_{\text{flows}} = 502$; 4b: $n_{\text{stock}} = 1,005$ and $n_{\text{flows}} = 996$) and then answered two questions (in random order) about when costs began to exceed income and when depletion was projected to happen.

The key addition in these studies was the addition of two self-reflection questions: first, a yes/no question that asked if participants thought Social Security would continue to collect payroll taxes after depletion; second, an open-ended question that asked participants about what they thought Social Security would do with the money if payroll taxes continued to be collected. These questions did not provide participants with any new information. Instead, the questions merely asked participants to consider for themselves the consequences of whatever they already knew about Social Security as a pay-as-you go system. Crucially, participants were randomly assigned to answer these questions as an intervention that came either before or after the questions about what would happen to benefits as a result of depletion.²¹ We refer to participants who saw the intervention before the questions about benefits as being in the treatment condition (4a: $n = 491$; 4b: $n = 998$) and those who saw the intervention after as being in the control condition (4a: $n = 510$; 4b: $n = 1,003$). Thus, these studies used a 2 (presentation: *stock* vs. *flow*) $\times$ 2 (intervention: *treatment* vs. *control*) design. The remainder of the surveys were the same as Experiments 1 and 2.²²

²⁰ See Supplemental Section 3 for a brief summary of an analysis looking at what participants write in at this question.

²¹ Of note, across all experiments (including these), this question already tells participants to "assum[e] the government does not take any action to increase the amount of income that Social Security collects based on tax revenues," which draws attention to the flows and could even imply that taxes will continue.

²² The only exception is that Experiment 4b did not include the question that asked participants to write a headline about the future of Social Security.

Results

We used regression analyses with contrast-coded predictors (−1, 1) to test the impact of presentation condition and intervention condition. As in previous studies, we also randomized the order of the first two objective knowledge questions about key dates, so all models included a contrast-coded predictor for order (−1, 1) that controls for this factor (we discuss findings related to this at the end of this section). Our models also include all two- and three-way interactions between factors. However, since none of these interactions are significant, we do not report on them below (see [Supplemental Section 4](#) for full model results). Following our preregistrations, we focus on stock-flow presentation effects only for the two date questions and both presentation and intervention effects for the questions about what happens to benefits.

The main results of these two experiments are summarized in [Table 5](#). As in previous studies, we did not see a significant main effect of presentation on accuracy in identifying when costs began to exceed income (4a: $b = 0.04$, $z = 0.61$, $p = .54$; 4b: $b = 0.01$, $z = 0.13$, $p = .89$). We did, however, replicate Experiment 1's finding that significantly more participants in the stock condition (4a: 79%; 4b: 78%) correctly identified the projected depletion date, compared to the flows condition (4a: 74%, $b = 0.17$, $z = 2.21$, $p = .027$; 4b: 69%, $b = 0.24$, $z = 4.71$, $p < .001$). The effect of intervention condition was nonsignificant for both of these dependent variables (4a income vs. costs: $b = 0.06$, $z = 0.85$, $p = .39$; 4b income vs. costs: $b = -0.01$, $z = -0.30$, $p = .76$; 4a depletion: $b = 0.03$, $z = 0.35$, $p = .73$; 4b depletion: $b = -0.02$, $z = -0.38$, $p = .70$), as expected given that everyone responded to these questions before the intervention.

Analyses focused on the zero-benefits beliefs answer (i.e., that benefits will go away completely) revealed that the main effect of presentation condition was not significant in Experiment 4a ($b = 0.04$, $z = 0.59$, $p = .56$), though results were directionally consistent with earlier studies. However, this effect was significant in Experiment 4b, replicating earlier findings that those in the stock condition were more likely to choose the answer option that benefits would go away completely (54% vs. 47%; $b = 0.14$, $z = 3.13$, $p = .002$).²³ The main effect of the intervention condition on answers to the question about what happens to benefits was significant and quite large in both studies. Participants in the treatment condition who were asked to reflect on what would happen to the inflow of payroll taxes before this question were significantly less likely to choose the zero-benefits beliefs answer option (4a: 43%; 4b: 39%) compared to those in the control condition who saw the intervention questions later (4a: 66%, $b = 0.48$, $z = 7.29$, $p < .001$; 4b: 62%, $b = 0.46$, $z = 10.04$, $p < .001$). A secondary, nonpreregistered analysis focused on the correct answer revealed that participants in the treatment condition who answered the intervention questions first were more likely than those in the control condition to indicate that benefits would still be paid out at a smaller amount (4a: 46% vs. 30%, $b = -0.36$, $z = -5.40$, $p < .001$; 4b: 48% vs. 29%, $b = -0.41$, $z = -8.74$, $p < .001$).

In both studies, the interactions between the presentation and intervention factors are not statistically significant (both $ps > .5$), which could suggest that the intervention that draws attention to the continuity of inflows is operating through a different channel than the graphs showing historical and projected flows. We may lack the statistical power to detect an interaction (especially with a binary

outcome variable; the simple effect is directionally a bit smaller in the intervention conditions). Additionally, the lack of an interaction need not be inconsistent with the flows presentations and the intervention both relying on the mechanism of accessibility of flows. The flows graphs may indirectly increase the broader tendency to reflect on inflows and outflows, and the intervention may increase the tendency to reflect on flows in a more direct and targeted manner. It is possible these effects may be additive.

Turning to the benefits amount question, the overall pattern of results was similar to previous studies. Among those who answered the question, the average amount written in was \$592 ($SD = \260) in Experiment 4a and \$606 ($SD = \291) in Experiment 4b; using imputed values based on our preregistered approach, the overall average was \$291 ($SD = \372) in Experiment 4a and \$334 ($SD = \400) in Experiment 4b. The main effect of stock-flow presentation was not significant in Experiment 4a, $b = -4.01$, $t(992) = -0.34$, $p = .73$, but was significant in Experiment 4b such that participants in the flows condition expected significantly higher monthly benefits ($M_{\text{flow}} = \$359$, $SD_{\text{flow}} = \$402$) than those in the stock condition, $M_{\text{stock}} = \$310$, $SD_{\text{stock}} = \$397$; $b = -21.63$, $t(1991) = -2.46$, $p = .014$. In both studies, the effect of the intervention condition was significant such that participants in the treatment condition who saw the intervention questions first estimated significantly higher benefits amounts after trust fund depletion (4a: $M_{\text{treatment}} = \$355$, $SD_{\text{treatment}} = \385 ; 4b: $M_{\text{treatment}} = \$404$, $SD_{\text{treatment}} = \403), compared to those in the control condition who saw the intervention questions later, 4a: $M_{\text{control}} = \$229$, $SD_{\text{control}} = \$348$, $b = -64.04$, $t(992) = -5.49$, $p < .001$; 4b: $M_{\text{control}} = \$265$, $SD_{\text{control}} = \$385$; $b = -68.35$, $t(1991) = -7.87$, $p < .001$. While still below the levels of what Social Security projections suggest could still be paid out (about \$750–\$800 based on the 2022 Trustees report), the mean amounts in the treatment condition are closer to realistic projections than in the flows conditions from prior studies.

Finally, as in our other experiments, we randomized the order in which we asked the first two objective knowledge questions about when costs began to exceed income and when depletion is projected to occur. Results in this study suggest a strong order effect on accuracy for both of these questions. Replicating our finding in Experiment 1, those who saw the question about the date when costs exceeded income first were less likely to answer correctly (4a: 47%; 4b: 50%) than those who saw that question second (4a: 65%, $b = 0.35$, $z = 5.49$, $p < .001$; 4b: 58%, $b = 0.16$, $z = 3.63$, $p < .001$). In Experiment 4a only, we also find a significant effect of order on

²³ The simple effect of stock versus flows in the control condition for Experiment 4a ($n = 510$) is directionally consistent with the stock-flow effect observed in other studies, with 68% in the stock condition choosing the wrong answer (vs. 65% in the flows condition). The simple effect in the treatment condition ($n = 491$) is directionally smaller, with 44% in the stock condition choosing the wrong answer (vs. 43% in the flows condition). The simple effects within the control and treatment conditions for Experiment 4b are in the expected direction and similar in magnitude to prior studies: in the control condition ($n = 1,003$), 66% chose the wrong answer in the stock condition (vs. 58% in the flows condition); in the treatment condition ($n = 998$), 42% chose the wrong answer in the stock condition (vs. 36% in the flows condition). These directionally smaller stock-flow effects in Experiment 4a, we argue, are consistent with sampling variability (particularly given our lack of a file drawer in estimating this effect across the studies in this article, the fact that the stock-flow factor was crossed with the “stronger” intervention manipulation in this study, and the larger stock-flow effect observed in the higher powered Experiment 4b).

Table 5*Summary Table of Results for All Preregistered Dependent Variables for Experiments 4a and 4b*

Measure ^a	Treatment		Control		Test of main effects	
	Stock	Flow	Stock	Flow	Treatment versus control	Stock versus flow
Experiment 4a						
Costs exceed income date (% correct) ^b	54%	56%	59%	55%	n.s.	n.s.
Depletion date (% correct) ^b	79%	73%	80%	74%	n.s.	$p = .027$
Zero-benefits beliefs (%) ^c	44%	43%	68%	65%	$p < .001$	n.s.
Benefits amount, $M (SD)$ ^d	\$361 (387)	\$349 (383)	\$214 (346)	\$242 (351)	$p < .001$	n.s.
Experiment 4b						
Costs exceed income date (% correct) ^b	52%	56%	56%	51%	n.s.	n.s.
Depletion date (% correct) ^b	78%	69%	78%	68%	n.s.	$p < .001$
Zero-benefits beliefs (%) ^c	42%	36%	66%	58%	$p < .001$	$p = .002$
Benefits amount, $M (SD)$ ^d	\$391 (405)	\$417 (401)	\$235 (374)	\$296 (393)	$p < .001$	$p = .014$

Note. n.s. = not statistically significant ($\alpha = .05$).

^aMissing answers due to a participant failing to answer a question are marked as missing but are included in the denominator for calculating the proportion of correct answers. These participants are excluded from regression analyses. ^bAs preregistered, answers are coded as correct if the participant entered a date within ± 1 year of the correct answer. ^cThe percentage of participants choosing this zero-benefits belief option is a key dependent variable across all studies. ^dFollowing our preregistration, responses above \$2,000 are excluded from calculations. The correct value is projected to be approximately \$750–\$800.

accuracy for the depletion date question such that participants who answered this question first were more likely to answer correctly (80%) than those who answered the question about income and costs first (73%; $b = 0.18$, $z = 2.38$, $p = .017$). This was not replicated in Experiment 4b ($b = 0.04$, $z = 0.76$, $p = .45$). However, there was a significant order effect in Experiment 4b on the question about what happens to benefits such that participants who saw the depletion date question second were less likely to choose the correct answer option about partial benefits being possible after depletion (41% vs. 36%; $b = -0.09$, $z = -1.98$, $p = .048$). We report the full results of all models in [Supplemental Section 4](#) and return to potential interpretations in the General Discussion section.

Discussion

In this pair of studies, we tested an intervention that drew participants' attention directly to the state of the inflows into the trust funds, asking them to consider whether Social Security would continue to collect payroll taxes after depletion and how that tax revenue would be used. Notably, a very large majority—90% of the sample in Experiment 4a and 89% in Experiment 4b—believed that Social Security would continue to receive these inflows, suggesting this specific aspect of Social Security policy is well understood. Importantly, however, those asked to reflect on this *before* answering a question about what would happen to benefits as a result of depletion were both less likely to think benefits would go away completely and more likely to choose the correct answer that partial benefits would still be paid out. Thus, this intervention, designed specifically to attenuate inflow neglect by prompting participants to reflect on the fact that the trust funds would continue to receive inflows (i.e., income from payroll taxes), was successful in combatting zero-benefits beliefs, the widely held misperception that Social Security benefits (i.e., outflows) will dry up when the trust funds do.

Experiment 5

Our final experiment served two main purposes. First, it enabled us to examine the generalizability of our findings that people neglect

inflows and hold zero-outflow beliefs and that this neglect is attenuated by flow-focused interventions, in a setting other than the depletion of the Social Security trust funds. Second, it enabled us to examine whether written descriptions focused on stock versus flows information would generate similar effects to those we observed with stock or flows graphs.

Materials and Procedure

We recruited a total of 1,091 unique participants from AMT to this study, 1,002 of whom completed it in full ($M_{\text{age}} = 41.80$, $SD_{\text{age}} = 12.11$; 41% male, 57% female, 1% nonbinary, 0.6% prefer not to say). The experiment used a factorial 2 (*modality*: graph vs. description) $\times$ 2 (*presentation*: stock vs. flow) $\times$ 2 (*intervention*: reflection prompt before vs. after) design. Participants first read a description of a fictional company that manufactures batteries for renewable energy storage. This description provided information on the current level of the company's inventory (i.e., the stock), its typical monthly production rates (i.e., inflows), and its typical shipments (i.e., outflows). Participants were told that demand for the company's product had increased recently, meaning the company was selling and shipping batteries faster than it could produce them and was starting to draw down on its inventory. At the end of the description, participants saw either a written blurb or a graph (corresponding to the *modality* factor, $n_{\text{graph}} = 497$, $n_{\text{written}} = 505$) that emphasized either the depleting inventory or the company's production and shipment rates (corresponding to the *presentation* factor, $n_{\text{stock}} = 505$, $n_{\text{flows}} = 497$); exact stimuli are presented in [Supplemental Section 6](#). After viewing this description, participants then answered two questions in a randomized order. The reflection question addressed what would happen to production if the inventory was depleted. The dependent variable addressed what would happen to shipments if the inventory was depleted. The order of these two questions varied depending on the *intervention* factor ($n_{\text{treatment}} = 506$, $n_{\text{control}} = 496$; those in the treatment condition answered the reflection question before the dependent variable, and those in the control condition answered the reflection question after the dependent variable). As preregistered, results below focus on the proportion of participants who indicated that

shipments would stop after the inventory is depleted (i.e., who hold “zero-outflow beliefs”).

Results

We used logistic regression analyses with contrast-coded predictors (−1, 1) to test the impact of the intervention, emphasis, and method factors. Our models also included all two- and three-way interactions between factors (we focus on the main effects below; see [Supplemental Section 4](#) for full model results).²⁴ Results across the stock versus flows and treatment versus control conditions are summarized in [Table 6](#) (see [Supplemental Section 5](#) for full results including the modality factor).

We were most interested in the main effect of the intervention factor—in other words, in this new setting, did asking participants to first reflect on what happens to inflows after depletion help them better reason through what would happen to outflows? Results demonstrated a similar pattern to what we observed in the Social Security setting: Those who answered the reflection prompt about inflows before indicating what would happen to outflows were significantly less likely to think that shipments would stop completely (23% vs. 41%; $b = -0.09$, $z = -6.36$, $p < .001$). As in prior studies, there was also a significant main effect of stock versus flows presentation, where those who saw descriptions or graphs that focused on the flows (vs. stock) were less likely to think that shipments would stop (30% vs. 35%; $b = -0.03$, $z = -2.23$, $p = .026$).²⁵ The main effect of modality was not significant ($b = -0.02$, $z = -1.12$, $p = .26$), and the difference between the stock and flows conditions was directionally consistent whether presented as a graph or as a written description, suggesting that emphasizing stock or flows information may have a similar impact on perceptions regardless of whether it is done through a graph or in words.

Discussion

In settings beyond the particular context of Social Security, this study was designed to explore the question of whether people neglect inflows when forecasting what will happen to outflows after the depletion of the stock. In the scenario tested here—regarding the inventory and shipments of a fictional company—a significant portion of participants still jumped to the conclusion that a depleted inventory meant an end to shipments of the product. As in earlier studies, graphs or descriptions that focused on the flow aspects of the situation decreased the proportion of participants with this misperception, and asking participants about inflows *before* asking about outflows decreased it even further. Despite notable differences between this study’s hypothetical scenario and the real-world Social Security scenario examined in earlier studies, the magnitude of the stock-flow and intervention effects were similar across studies.

General Discussion

In this article, we explore the novel stock-flow reasoning error of inflow neglect. We do so both in the concrete, important setting of the impending depletion of the Social Security trust funds and in an unrelated, more controlled inventory management scenario. In both contexts, we find that substantial proportions of participants erroneously indicate that outflows will cease after the depletion of the stock. We argue that one reason for these problematic interpretations is that people neglect to consider the continuation of system inflows and,

especially when communications focus on the impending depletion of the stock, jump to the conclusion that outflows will cease in the future.

Across five preregistered experiments, we tested whether differences in inferences regarding the Social Security trust funds arose based on reasoning about stocks versus flows. Results from Experiments 1, 3, 4a, and 4b suggest that presenting the information as a stock leads to higher accuracy regarding *when* depletion of the trust funds will occur (this effect was not significant in Experiment 2). This held despite the fact that all participants read a description that included the correct depletion date. But as shown in our stimuli (see [Figures 1 and 2](#)), the stock graph emphasized that the balance would be running down to zero around 2035.

Our main results, however, center on inferences about what happens to outflows after depletion. Across all of our experiments, relative to those in the flows conditions, more respondents in the stock conditions thought that outflows would stop completely once the stock was depleted. Those exposed to alternative flows presentations—one version explicitly graphing payable benefits (in Experiment 2) and another integrating inflows and outflows into a single net flows presentation (in Experiment 3)—responded similarly to those who saw graphs that simply plotted the inputs and outputs, suggesting that drawing participant attention to the flows at all may enhance reasoning about what happens to future outflows. While these effects are modest in magnitude, they are notable given that all conditions included reminders about the stock running out. These effects are also plausibly substantively important, especially for the Social Security setting, where media coverage tends to focus on the impending depletion date.

Flows presentations, however, are not a panacea: More than half of those in the flows conditions across Experiments 1–3 still held zero-outflow beliefs. To this end, Experiments 4a, 4b, and 5 tested an intervention aimed at further reducing the misconception that outflows stop completely when the stock becomes depleted and provided supportive process evidence about the stock-flow reasoning error underlying these beliefs (i.e., inflow neglect). This intervention was simple: We asked targeted questions that asked participants to explicitly consider whether inflows would continue postdepletion. Despite its simplicity, asking these questions before (vs. after) the questions about what happens to outflows resulted in a large reduction in the number of people who thought outflows would cease completely, both in the Social Security setting and in an unrelated manufacturing scenario.

Although there may be ways to improve understanding even further, the results of these studies point to this kind of targeted reflection being a promising technique for harnessing beliefs about the continuity of

²⁴ Across all models, all two-way and three-way interactions are nonsignificant, save for one exception. The interaction between the presentation and intervention factor was significant for the outcome of choosing the correct answer about future outflows ($p = .026$), such that in the treatment condition, those who saw flows presentations are more likely to choose the correct answer than those who saw stock presentations, but in the control condition, those who saw flows presentations are less likely to choose the correct answer. There is no such interaction when the outcome is choosing the incorrect answer. It seems that in the control condition, those who see the flows presentation are shifting more into the “shipments continue as is” category.

²⁵ Consistent with our proposed explanation—that increasing the accessibility of inflows helps participants reason about outflows—those who saw descriptions and/or graphs that emphasized the flows (vs. the stock) were more likely to correctly identify that production will continue after depletion (90% vs. 83%; $b = 0.04$, $z = 3.51$, $p < .001$). Across all conditions (i.e., when including all condition contrast codes), those who correctly answer the reflection question were also less likely to think that outflows would stop completely (27% vs. 68%; $b = -0.40$, $z = -9.71$, $p < .001$).

Table 6
Summary Table of Results for Experiment 5

What happens to shipments question (% choosing each option) ^a	Treatment		Control		Test of main effect	
	Stock	Flow	Stock	Flow	Treatment versus control	Stock versus flow
Shipments stop completely ^b	27%	19%	44%	39%	$p < .001$	$p = .026$
Shipments slow down ^c	61%	68%	50%	43%	$p < .001$	n.s.
Shipments continue as is	10%	11%	4%	10%		
Shipments speed up	1%	3%	3%	7%		

Note. $N = 1,002$. In this experiment, there was a significant main effect of intervention (where those in the treatment condition who were asked to reflect on inflows first were less likely to choose the incorrect answer that shipments stop completely) and a significant main effect of emphasis (where those shown flows [vs. stock] information are less likely to choose the incorrect answer that shipments stop completely). The main effect of method (graph vs. written emphasis) is not significant. n.s. = not statistically significant ($\alpha = .05$).

^aMissing answers due to a participant failing to answer a question are marked as missing but are included in the denominator for calculating the proportion of correct answers. These participants are excluded from regression analyses. ^bWhether or not participants choose this answer option is our key dependent variable. ^cThis is the correct answer.

inflows to reinforce expectations for the continuity of outflows. It is also informative about the psychological process by indicating that participants may be capable of the necessary stock-flow reasoning when prompted to reflect upon how the parts of the system fit together. Note that the intervention question did not provide participants with any new information; instead, it seems to have generated an “aha” moment where the acknowledgment of continued inflows is used to infer that outflows will continue too. This approach may provide a promising avenue for improving qualitative stock-flow reasoning problems more generally, which have previously proven stubbornly resistant to accuracy interventions (e.g., Cronin et al., 2009).

Unexpectedly, in Experiments 1, 3, 4a, and 4b (though not Experiment 2), we also found that participants were more likely to correctly report the year in which Social Security’s outflows began to exceed inflows when they did so after reporting when the trust fund would be depleted. Revealingly, among participants who did not enter the correct date for when costs started exceeding income, a majority entered the depletion date instead. This suggests that when encountering the question regarding flows after the question regarding stocks, respondents were more likely to properly distinguish between the two representations. In contrast, when encountering the question regarding flows first, they may not have distinguished the two metrics and instead answered with respect to the stock, reinforcing the idea that the depletion of the stock was the most salient aspect of the system as presented. Further investigation of answers to these questions about key dates and answers about what will happen to benefits revealed that participants who answered the depletion date question correctly but the income-versus-cost date incorrectly (which may reflect a relative focus on the stock) were the *most* likely to have zero-benefits beliefs while those who answered the depletion date incorrectly but the income-versus-cost date correctly (which may reflect a relative focus on the flows) were among the *least* likely to have zero-benefits beliefs. Taken together, this evidence is consistent with participants’ answers reflecting an overall focus on the stock versus the flows (which may or may not be completely accurate for understanding all the pieces of the system).

This work holds both theoretical and practical importance. Theoretically, the Social Security setting provides an opportunity to test the effect of stock versus flows presentations on a new type of error (i.e., inflow neglect) and a novel outcome (i.e., a projection about what will happen to a component of the system in the future—in this case, the

outflows or benefits payments). Our final experiment highlights how the insights from our Social Security experiments also extend beyond that specific context, pointing to inflow neglect as a broader reasoning error in stock-flow situations involving impending depletion. From a practical perspective, our results suggest that communicators attempting to overcome misconceptions about future outflows should consider presenting information in terms of flows or prompting readers to consider their existing expectations for the system’s future inflows. This may be especially important for promoting public understanding of the financial future of Social Security, which has wide-ranging consequences for prospective retirees.

Constraints on Generality

This investigation has its limitations. First, our studies were conducted solely on AMT. Findings using online convenience samples like AMT are strongly correlated with findings using (typically much more expensive) probability samples (Coppock et al., 2018; Mullinix et al., 2015; Peyton et al., 2021; Snowberg & Yariv, 2021), especially when using screeners such as the CloudResearch approved pool we used (Litman et al., 2017; Peer et al., 2021). Second, most of our studies were focused on a policy outcome directly relevant for prospective retirees in the United States, though we expect (and find in at least one example in Experiment 5) that the stock-flow reasoning insights resulting from this article will apply to other contexts involving accumulation. Third, our studies did not include incentive-compatible outcomes or consequential choices, and so it is possible that incentives could have enhanced accuracy (although unless they also would have reduced the difference between conditions, they would not pose a threat to our estimate of the effect of stock vs. flows presentation or the effect of our intervention; prior research suggests enhanced motivation is not sufficient; e.g., Cronin et al., 2009). Finally, even though we attempted to generate stimuli that closely resembled material used by the Social Security Administration, our materials were written to be more accessible and understandable than the status quo. It is possible that differences between conditions would be muted were we to use actual SSA materials.

Future Directions and Conclusion

Despite these limitations, the results are promising and offer opportunities for future research. Specifically, more work is needed to

better understand how to further improve de-biasing attempts. We reason that it may be difficult to alter perceptions about the trust funds precisely because the stock framing of the trust funds is so strongly ingrained. Indeed, media headlines themselves may reinforce a focus on stocks rather than flows (cf. Jerit & Barabas, 2006). Further, because stock-flow reasoning is so difficult, future interventions may need to be more involved and/or occur over multiple time points rather than the single-shot interventions we employed in these studies (e.g., the further-reinforced “enhanced flow” intervention in Experiment 2 or the reflection intervention in Experiments 4a, 4b, and 5). In this article, we have explored inflow neglect in two distinct settings facing impending resource depletion—the important, real-world setting of Social Security and a controlled, hypothetical setting involving inventory management—and we hope that future research will examine additional contexts in which people may be prone to inflow neglect, such as the drawdown of greenhouse gases from the atmosphere, the depletion of important inventories or resources (e.g., water reservoirs, blood banks), or the budgets of organizations and individuals.

Specific to the retirement context, our experiments highlight how stock-flow reasoning might contribute to prospective retirees’ misunderstanding of the future of the Social Security system, but beliefs about the end of future benefits may have additional root causes. In exploratory analyses, we found that across Experiments 1–4b, participants who had attained a college degree or higher were less likely to hold zero-benefits beliefs (vs. those who had not attained a college degree). However, the effect of flows presentations and manipulations on zero-benefits beliefs did not consistently depend on education.²⁶ Thus, it may be that more educated participants simply knew more about the system at baseline or could understand the description of the trust funds more easily, but additional research is required to fully understand this relationship. Beyond individual differences, it could also be interesting to explore interactions with the influence of the political climate and other uncertainties that might affect perceptions of the Trustees’ projections. Finally, we focused on objective understanding of the system, which is an important outcome in and of itself—but objective understanding is also important because it may affect downstream consequences such as claiming age or savings behavior (see, e.g., Delavande & Rohwedder, 2011; Quinby & Wettstein, 2021). Future work may benefit from directly studying these downstream consequences, as well as other subjective outcomes (e.g., concern about the situation) and support for changes to SSA policies regarding taxes and benefits to address the long-term funding shortfall.

More generally, we have argued that inflow neglect arises, at least in part, from low accessibility of flows-related information in the face of impending depletion. That said, the larger theoretical question of *why* individuals tend to focus on the depletion of the stock, rather than on the continuation of the flows, is a ripe area for further inquiry. Prior work on how individuals preemptively react to anticipated low inventories by building up reserves or buffer stocks suggests that concerns about loss aversion and preservation of options to cover emergencies may be important drivers (e.g., S. Johnson et al., 1987; Sharif & Shu, 2017; Shin & Ariely, 2004; Shu, 2008; Walker et al., 2018). Individual differences in risk aversion, loss aversion, and sensitivity to opportunity costs are all possible moderators of these concerns (Greenberg & Spiller, 2016; Spiller, 2011), and additional factors such as education or math ability may moderate susceptibility to inflow neglect in stock-flow settings, as noted above. Finally, a range of cognitive shortcuts may also contribute to or exacerbate inflow neglect errors, such as the correlation heuristic described in the

introduction (Cronin et al., 2009; Fansher et al., 2025; Fischer & Gonzalez, 2016) or the headwinds–tailwinds asymmetry in which people pay more attention to obstacles that hinder progress compared to forces that aid goal completion (Davidai & Gilovich, 2016).

Governments, companies, and households are constantly making decisions involving the accumulation and decumulation of important assets, and forecasts of future flows are a critical part of this calculus. It is important that decision-makers fully understand these situations in order to make informed decisions—both when managing their own resources and when asked to weigh in on matters of public policy. This work identifies inflow neglect as a novel stock-flow reasoning error with important consequences for such forecasts, contributing to a deeper understanding of how people evaluate complex system dynamics and pointing to how communications about these dynamics can influence public understanding and discourse.

²⁶ There are no significant interactions between education and condition in Experiments 1, 3, 4a, and 4b. The one exception is Experiment 2, where there is a marginally significant interaction ($b = 0.50$, $z = 1.88$, $p = .06$), such that the difference in zero-benefits beliefs between the stock versus plain flows condition was larger among those who have attained a college degree or higher. One possible post hoc explanation for this one-off finding is that the graphs used in this study were particularly complex; however, further research is needed replicate and probe these exploratory results.

References

- The Board of Trustees, Federal Old-Age and Survivors Insurance and Federal Disability Insurance Trust Funds. (2025). *The 2025 annual report of the board of trustees of the federal old-age and survivors insurance and federal disability insurance trust funds*. <https://web.archive.org/web/20250315051746/https://www.ssa.gov/OACT/TR/2025/tr2025.pdf>
- Booth Sweeney, L., & Sterman, J. D. (2000). Bathtub dynamics: Initial results of a systems thinking inventory. *System Dynamics Review*, 16(4), 249–286. <https://doi.org/10.1002/sdr.198>
- Boz-Yilmaz, H. Z., & Boduroglu, A. (2024). Understanding patterns of accumulation: Improving forecast-based decisions via nudging. *Memory & Cognition*, 52(5), 1033–1047. <https://doi.org/10.3758/s13421-024-01519-6>
- Brunstein, A., Gonzalez, C., & Kanter, S. (2010). Effects of domain experience in the stock–flow failure. *System Dynamics Review*, 26(4), 347–354. <https://doi.org/10.1002/sdr.448>
- Coppock, A., Leeper, T. J., & Mullinix, K. J. (2018). Generalizability of heterogeneous treatment effect estimates across samples. *Proceedings of the National Academy of Sciences of the United States of America*, 115(49), 12441–12446. <https://doi.org/10.1073/pnas.1808083115>
- Cronin, M. A., Gonzalez, C., & Sterman, J. D. (2009). Why don’t well-educated adults understand accumulation? A challenge to researchers, educators, and citizens. *Organizational Behavior and Human Decision Processes*, 108(1), 116–130. <https://doi.org/10.1016/j.obhdp.2008.03.003>
- Davidai, S., & Gilovich, T. (2016). The headwinds/tailwinds asymmetry: An availability bias in assessments of barriers and blessings. *Journal of Personality and Social Psychology*, 111(6), 835–851. <https://doi.org/10.1037/pspa0000066>
- Delavande, A., & Rohwedder, S. (2011). Individuals’ uncertainty about future Social Security benefits and portfolio choice. *Journal of Applied Econometrics*, 26(3), 498–519. <https://doi.org/10.1002/jae.1235>
- Fansher, M., Lalwani, P., Adkins, T. J., Zhang, H., Quirk, M., Carlson, M., Boduroglu, A., Lewis, R. L., Jonides, J., & Shah, P. (2025). A brief intervention to improve reasoning about accumulation. *Journal of Experimental Psychology: Applied*, 31(2), 99–125. <https://doi.org/10.1037/xap0000532>

- Feldman, J. M., & Lynch, J. G. (1988). Self-generated validity and other effects of measurement on belief, attitude, intention, and behavior. *Journal of Applied Psychology*, 73(3), 421–435. <https://doi.org/10.1037/0021-9010.73.3.421>
- Fischer, H., & Gonzalez, C. (2016). Making sense of dynamic systems: How our understanding of stocks and flows depends on a global perspective. *Cognitive Science*, 40(2), 496–512. <https://doi.org/10.1111/cogs.12239>
- Goldstein, D. G., Herschfield, H. E., & Benartzi, S. (2016). The illusion of wealth and its reversal. *Journal of Marketing Research*, 53(5), 804–813. <https://doi.org/10.1509/jmr.14.0652>
- Greenberg, A. E., & Spiller, S. A. (2016). Opportunity cost neglect attenuates the effect of choices on preferences. *Psychological Science*, 27(1), 103–113. <https://doi.org/10.1177/0956797615608267>
- Harman, J. L., Weinhardt, J. M., Beck, J. W., & Mai, I. (2021). Interpreting time-series COVID data: Reasoning biases, risk perception, and support for public health measures. *Scientific Reports*, 11(1), Article 15585. <https://doi.org/10.1038/s41598-021-95134-z>
- Hoffrage, U., Lindsey, S., Hertwig, R., & Gigerenzer, G. (2000). Communicating statistical information. *Science*, 290(5500), 2261–2262. <https://doi.org/10.1126/science.290.5500.2261>
- Holahan, W. L., & Schug, M. C. (2000). A simple exposition of the Social Security trust fund. *The Journal of Economic Education*, 31(4), 340–348. <https://doi.org/10.1080/00220480009596451>
- Hutchinson, J. W., Alba, J. W., & Eisenstein, E. M. (2010). Heuristics and biases in data-based decision making: Effects of experience, training, and graphical data displays. *Journal of Marketing Research*, 47(4), 627–642. <https://doi.org/10.1509/jmkr.47.4.627>
- Jerit, J., & Barabas, J. (2006). Bankrupt rhetoric: How misleading information affects knowledge about Social Security. *Public Opinion Quarterly*, 70(3), 278–303. <https://doi.org/10.1093/poq/nfl010>
- Johnson, E. J., Häubl, G., & Keinan, A. (2007). Aspects of endowment: A query theory of value construction. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33(3), 461–474. <https://doi.org/10.1037/0278-7393.33.3.461>
- Johnson, S., Kotlikoff, L. J., & Samuelson, W. (1987). *Can people compute? An experimental test of the life cycle consumption model* (NBER Working Paper No. 2183). <https://www.nber.org/papers/w2183>
- Kiely, E. (2022, March 29). *Examining Rick Scott's claim that medicare, Social Security will soon go "bankrupt"*. Factcheck.org. <https://www.factcheck.org/2022/03/examining-rick-scotts-claim-that-medicare-social-security-will-soon-go-bankrupt/>
- Litman, L., Robinson, J., & Abberbock, T. (2017). TurkPrime.com: A versatile crowdsourcing data acquisition platform for the behavioral sciences. *Behavior Research Methods*, 49(2), 433–442. <https://doi.org/10.3758/s13428-016-0727-z>
- Luttmer, E. F. P., & Samwick, A. A. (2018). The welfare cost of perceived policy uncertainty: Evidence from Social Security. *American Economic Review*, 108(2), 275–307. <https://doi.org/10.1257/aer.20151703>
- Mullinix, K. J., Leeper, T. J., Druckman, J. N., & Freese, J. (2015). The generalizability of survey experiments. *Journal of Experimental Political Science*, 2(2), 109–138. <https://doi.org/10.1017/xps.2015.19>
- Newell, B. R., Kary, A., Moore, C., & Gonzalez, C. (2016). Managing the budget: Stock-flow reasoning and the CO₂ accumulation problem. *Topics in Cognitive Science*, 8(1), 138–159. <https://doi.org/10.1111/tops.12176>
- Padilla, L., Hosseinpour, H., Fygenon, R., Howell, J., Chunara, R., & Bertini, E. (2022). Impact of COVID-19 forecast visualizations on pandemic risk perceptions. *Scientific Reports*, 12(1), Article 1014. <https://doi.org/10.1038/s41598-022-05353-1>
- Paul, T. (2022, July 8). *When will Social Security run out? Here's what could happen to your benefits*. CNBC. <https://www.cnbc.com/select/social-security-run-out-heres-what-you-need-to-know/>
- Peer, E., Rothschild, D., Gordon, A., Evernden, Z., & Damer, E. (2021). Data quality of platforms and panels for online behavioral research. *Behavior Research Methods*, 54(4), 1643–1662. <https://doi.org/10.3758/s13428-021-01694-3>
- Peyton, K., Huber, G. A., & Coppock, A. (2021). The generalizability of online experiments conducted during the COVID-19 pandemic. *Journal of Experimental Political Science*, 9(3), 379–394. <https://doi.org/10.1017/xps.2021.17>
- Quinby, L., & Wettstein, G. (2021). *Does media coverage of the Social Security trust fund affect claiming, saving, and benefit expectations?* SSRN. <https://doi.org/10.2139/ssrn.4669070>
- Reinholtz, N., Maglio, S. J., & Spiller, S. A. (2021). Stocks, flows, and risk response to pandemic data. *Journal of Experimental Psychology: Applied*, 27(4), 657–668. <https://doi.org/10.1037/xap0000395>
- Sharif, M. A., & Shu, S. B. (2017). The benefits of emergency reserves: Greater preference and persistence for goals that have slack with a cost. *Journal of Marketing Research*, 54(3), 495–509. <https://doi.org/10.1509/jmr.15.0231>
- Shin, J., & Ariely, D. (2004). Keeping doors open: The effect of unavailability on incentives to keep options viable. *Management Science*, 50(5), 575–586. <https://doi.org/10.1287/mnsc.1030.0148>
- Shu, S. B. (2008). Future-biased search: The quest for the ideal. *Journal of Behavioral Decision Making*, 21(4), 352–377. <https://doi.org/10.1002/bdm.593>
- Smyth, A. H. (1907). *The writings of Benjamin Franklin (Vol. X) 1789–1790*. Macmillan.
- Snowberg, E., & Yariv, L. (2021). Testing the waters: Behavior across participant pools. *American Economic Review*, 111(2), 687–719. <https://doi.org/10.1257/aer.20181065>
- Spiller, S. A. (2011). Opportunity cost consideration. *Journal of Consumer Research*, 38(4), 595–610. <https://doi.org/10.1086/660045>
- Spiller, S. A., Reinholtz, N., & Maglio, S. J. (2020). Judgments based on stocks and flows: Different presentations of the same data can lead to opposing inferences. *Management Science*, 66(5), 2213–2231. <https://doi.org/10.1287/mnsc.2019.3284>
- Sterman, J. D. (2008). Risk communication on climate: Mental models and mass balance. *Science*, 322(5901), 532–533. <https://doi.org/10.1126/science.1162574>
- Sterman, J. D., & Booth Sweeney, L. (2007). Understanding public complacency about climate change: Adults' mental models of climate change violate conservation of matter. *Climatic Change*, 80(3), 213–238. <https://doi.org/10.1007/s10584-006-9107-5>
- Turner, J. A., & Rajnes, D. (2021). Workers' expectations about their future Social Security benefits: How realistic are they? *Social Security Bulletin*, 81(4), 1–17. <https://www.ssa.gov/policy/docs/ssb/v81n4/v81n4p1.html>
- Villanova, D. (2022). Linear biases and pandemic communications. *Medical Decision Making*, 42(6), 765–775. <https://doi.org/10.1177/0272989x221107907>
- Walker, J., Risen, J. L., Gilovich, T., & Thaler, R. (2018). Sudden-death aversion: Avoiding superior options because they feel riskier. *Journal of Personality and Social Psychology*, 115(3), 363–378. <https://doi.org/10.1037/pspa0000106>
- Weber, E. U., Johnson, E. J., Milch, K. F., Chang, H., Brodscholl, J. C., & Goldstein, D. G. (2007). Asymmetric discounting in intertemporal choice: A query-theory account. *Psychological Science*, 18(6), 516–523. <https://doi.org/10.1111/j.1467-9280.2007.01932.x>
- Werschul, B. (2021, September 7). *Social Security expert details the "worst case" scenario—And why millennials are most vulnerable*. Yahoo! Finance. <https://finance.yahoo.com/news/social-security-biggest-risk-millennials-most-vulnerable-153719368.html>

Received August 20, 2024

Revision received August 15, 2025

Accepted October 29, 2025 ■

Reproduced with permission of copyright owner. Further reproduction
prohibited without permission.