JDDM
Journal of Dynamic Decision Making

Open Access

Original Research

Valence Matters in Judgments of Stock Accumulation in Blood Glucose Control and Other Global Problems

Cleotilde Gonzalez[1], **Maria-Isabel Sanchez-Segura**[2], **German-Lenin Dugarte-Peña**[2], **Fuensanta Medina-Dominguez**[2]

[1]Carnegie Mellon University, USA and [2] Universidad Carlos III de Madrid, Spain

Stock-flow failure is a reasoning error in dynamic systems that has great societal relevance: people misjudge a level of accumulation (i.e., stock) considering the information on flows that increase (i.e., inflow) or decrease (i.e., outflow) over time. Many interventions, including the use of analogies and graphical manipulations, to counteract this failure and help people integrate the flow information have been tested with little or no success. We suggest that this error relates to the *valence* of a problem: the framing of the inflow or outflow direction as "good" or "bad" is associated with the direction of its accumulation over time. To explore the effects of *valence* on accumulation judgments, we employ a scenario of a common health problem: blood glucose control through sugar consumption and insulin flows. We also investigate improvements of performance in a second scenario that result after viewing a video of a dynamic system demonstrating the effects of the correct accumulation trend. We discuss the implications of our findings for the blood glucose example and other global problems.

Keywords: Stock-Flow Failure, Correlation Heuristic, System Behaviour, Valence Effect

Dynamic systems are present in many daily life situations. Generally, people rarely even notice that they have the property of changing continuously over time (that is, their dynamics). Possible examples of dynamic systems are: population growth, learning processes in supply chains, inventory management, and savings and debt accumulation.

While dynamic systems are very common in real life, research into complex problem solving and dynamic decision making have shown that people have great difficulty learning and making decisions with respect to dynamic tasks, even after lengthy practice over long or unlimited time periods and with performance incentives (Diehl & Sterman, 1995; Frensch & Funke, 1995). To understand these difficulties, researchers have relied on simple abstractions of complex systems in order to study the basic elements of every dynamic system (Booth Sweeney & Sterman, 2000; Cronin & Gonzalez, 2007): stocks and flows (inflow, outflow). Inflow is the kind of flow that adds units to the stock, and outflow is the kind of flow that subtracts units from the stock. A stock increase occurs only if the inflow rate is higher than the outflow rate, and a stock decrease occurs if the inflow rate is lower than the outflow rate. This relationship between flows and stocks over time defines a behavioural pattern that is often represented in an x-y plot. In fact, graphs are commonly used to illustrate the behavioural patterns of dynamic systems, where time is plotted on the x-axis. For example, the world population growth is often represented on an x-y plot, with the year on the x-axis and the accumulation of billions of people or the annual growth rate (i.e., inflow minus outflow) on the y-axis.

Evidence from laboratory experiments using graphical methods and simple representations of dynamic tasks over the last decade suggests that people generally misunderstand the basic behavioural patterns of dynamic systems: the stock-flow failure (SF failure). The SF failure is evidence that people misinterpret the accumulation of a quantity ("stock") and often reason that there should be a direct relationship between the accumulation and the direction of the rates of change that it accumulates (i.e., the flows: inflow or outflow) (i.e., they should be positively correlated) (Cronin, Gonzalez, & Sterman, 2009). This is equivalent, for example, to deciding that the world population should decrease if the annual growth rate of the world population decreases. This is, of course, incorrect, and the world population will continue to increase as long as the net rate of growth is greater than zero (e.g., more people are born than die).

Booth-Sweeney and Sterman (2000) were the first to show that, when asked to plot the trajectory of an accumulating stock, people often draw a curve that matches the pattern of the flows. This phenomenon, later termed the *correlation heuristic*, was found to be very resistant to a wide range of interventions (actions designed to bring about changes in people) altering motivation, context, and mode of information representation (Cronin, Gonzalez, & Sterman, 2009). This fundamental misunderstanding of how system inflows and outflows accumulate over time contributes to a wide range of real-world problems at the personal, organizational and global levels, such as maintaining a healthy body weight (Abdel-Hamid et al., 2014), reducing atmospheric CO_2 (Dutt & Gonzalez, 2012a; Dutt & Gonzalez, 2012b; Newell, Kary, Moore, & Gonzalez, 2016; Sterman, 2008; Sterman & Booth Sweeney, 2002), and balancing budgets (Booth Sweeney & Sterman, 2000; Newell et al., 2016). One often-cited example of the correlation heuristic is the misunderstanding of how the concentration

Corresponding author: Cleotilde Gonzalez, 5000 Forbes Ave, Porter Hall 208, Pittsburgh, PA, 15213, USA. e-mail: coty@cmu.edu

of CO_2 in the atmosphere (the stock) relates to the CO_2 emitted into the atmosphere (the inflow) and the CO_2 absorbed via natural processes (the outflow). If the CO_2 concentration rises gradually and then stabilizes, people reason that the CO_2 inflow pattern (emissions) will increase similarly when absorptions are lower than emissions and stable over time. In reality, CO_2 emissions need to *decrease* over time and converge with the level of absorptions in order to achieve stabilization. This simple but very impactful misunderstanding often leads to erroneous decision making and policies when addressing climate change (Dutt & Gonzalez, 2012a; Dutt & Gonzalez, 2012b; Sterman, 2008).

While the potential policy implications of SF failure in applied societal problems like climate change may be clear, an understanding of the causes and possible ways to combat SF failure are often missing. In this research, we experimentally investigate a *valence* effect, first suggested by Newell and colleagues (2016), and we also explore the potential of a video demonstration as an intervention for the improvement of performance in judgments of stock and flows. As discussed in the literature review below, the valence hypothesis suggests that the way a problem is framed (i.e., the "goodness" or "badness" of a situation) determines the applicability of the correlation heuristic, and thus SF problem behaviour. To test this hypothesis, we use a relevant societal health stock-flow problem as an example: blood glucose control. In this problem, people often fail to understand how sugar accumulates in the bloodstream over time as a function of sugar intake (via food) and sugar absorption (via insulin or exercise). Below we also review the literature suggesting that video interventions have the potential to improve performance in SF tasks.

The valence effect

Newell et al. (2016) first hinted at a possible "valence" explanation for SF failure that emerges from the framing of a problem, the accumulation trend, and the direction of the flows modifying a stock. In an effort to improve performance in Sterman and Booth-Sweeney's (2007) CO_2 task, Newell and colleagues (2016) used financial debt management and savings management isomorphs of the CO_2 task. Applying the same methods as Sterman and Booth-Sweeney (2007), Newell et al. (2016) showed participants a graph of a stock (savings, debt, or CO_2) that increased over time and stabilized over a future time period, as well as a graph of the corresponding flows (expenses and earnings for the debt and savings contexts; emissions and absorptions for the CO_2 context). However, one of the flow curves in the flows graph was incomplete (whereas the other one was stationary at a lower level), and participants were asked to provide the value that the incomplete flow curve should equal at the end of the time period for the stock to stabilize as shown. In all contexts, the correct answer entailed *decreasing* the value of the incomplete flow to match the value of the other (stationary and lower) flow.

Newell et al. (2016) suspected that performance would be better for the debt/savings isomorphs than for the CO_2 task because people are more familiar with the debt and savings contexts. However, their results suggest that familiarity with the context alone does not support reasoning. Familiarity with the debt and savings isomorphs was not enough for participants to do well in these tasks, supporting the conclusions reached by Brunstein, Gonzalez, & Kanter (2010) and others regarding context familiarity. Using SF problems in the medical domain and comparing the performance of medical and general students, Brunstein et al. (2010) concluded that domain experience "*is not a strong indicator for overcoming the SF failure*" (page 352). The SF failure remained latent regardless of participant "expertise", taking into account that expertise itself is a complex cognitive phenomenon that depends on the domain in which it is considered and on the view or perspective used for its interpretation (Sternberg, 1995).

Instead, Newell et al. (2016) found that participants were more accurate when the problem was framed as increasing "debt" (and participants needed to decrease the "spending" to match the "earnings") than when the problem was framed as increasing "savings" (and participants needed to decrease "earnings" to match "spending"), even though the two problems were equivalent. Newell et al. (2016) suggested problem *valence* as an explanation, reasoning that, to do well in the SF task, the valence of the user-controlled flow should match the direction of the correct solution of the problem. Their conjecture was that people were more accurate with the debt framing, because participants reasoned that debt is "bad" and spending should *decrease*, whereas they were less accurate with the savings framing because savings are "good" and earnings should *increase*. In the former case, the valence of the user-controlled flow matched the direction of the correct solution (i.e., spending is "bad"), leading participants to correctly decrease the spending flow, overriding the correlation heuristic. In the latter case, the valence of the user-controlled flow opposed the direction of the correct solution (i.e., savings are "good"), leading participants to incorrectly increase the earnings flow in accordance with the correlation heuristic.

As Newell et al. (2016) had not originally aimed at testing the *valence hypothesis*, they used only increasing accumulation scenarios (framed as debt or savings) where the correct response required participants to decrease the user-controlled flow (spending or earnings). Therefore, we do not know the full extent to which the valence effect can counteract the correlation heuristic. For example, the performance improvement in the "debt" scenario in Newell et al. (2016) may have been fortuitous rather than the result of an actual understanding of stock-flow relationships, because the valence of the user-controlled "debt" flow matched the direction of the correct solution.

In this paper, we test the valence hypothesis by controlling for (1) the behaviour of the stock (increasing or decreasing), and (2) whether the user-controlled inflow or outflow matches or opposes the direction of the correct solution. We use a blood glucose management example because blood glucose increases and decreases can actually be both "good" or "bad"; high blood glucose (hyperglycemia) and low blood glucose (hypoglycemia) are major health concerns; established methods for controlling and stabilizing blood glucose entail both increasing and decreasing sugar consumption (inflow) and insulin (outflow); and many people fail to understand how sugar accumulates in the bloodstream over time as a function of sugar intake (via food) and sugar absorption (via insulin or exercise). Thus, blood glucose management is an ideal context to test the valence hypothesis.

According to the correlation heuristic, we hypothesize that, regardless of the direction of the stock (increasing or decreasing accumulation), when the valence of the user-controlled flow agrees with the direction of the correct response, performance will be better than when the user-controlled flow opposes the direction of the correct response. Furthermore, when the user-controlled flow opposes the direction of the correct response, we expect the valence of the user-controlled flow (i.e., soda or insulin intake) to influence the effect of the correlation heuristic. For example, when the valence of the user-controlled flow is "good" (i.e., decrease soda consumption), performance would be better than when the valence of the user-controlled flow is "bad" (i.e., increase soda consumption).

Potential of video demonstrations to improve SF failure

As mentioned above, people's ability to reason about stocks and flows has traditionally been tested by showing a graph of the system inflow and outflow and asking them to plot the resulting stock curve (Booth-Sweeney & Sterman, 2000; Sterman & Booth Sweeney, 2002). However, SF failure cannot entirely be explained by a failure to interpret graphs (see, for example, Cronin, Gonzalez, & Sterman, 2009; Fischer, Degen, & Funke, 2015). Also, SF failure persists even when the problem is posed as the physical, naturalistic task of pouring water through a funnel into a beaker to meet a target goal (Strohhecker & Größler, 2015). Key studies have attempted to improve people's understanding of accumulation focusing on the use of analogies (Gonzalez & Wong, 2012; Newell et al., 2016), priming (Fischer & Gonzalez, 2016), and alternative SF problem presentation styles (Fischer, Degen, & Funke, 2015; Fischer & Gonzalez, 2016), with limited success. Importantly, research suggests that interventions involving simulations and learning tools may be more effective in reducing SF failure (Dutt & Gonzalez, 2012a; Moxnes & Saysel, 2009; Sterman et al., 2012). A possible explanation for the improvement is that interactive simulations are experiential rather than descriptive. For example, Dutt and Gonzalez (2012b) exposed participants to descriptive and experiential CO_2 stabilization task conditions. The descriptive version emulated the task designed by Sterman and Booth-Sweeney (2007), whereas participants used a dynamic climate change simulator to perform the same CO_2 stabilization task in the simulated version. They found that people's misconceptions decreased significantly when they practiced the task using the simulation. This finding highlights the potential for using experience-based simulation tools to improve the understanding of the dynamics of climate change.

In this study, we employ a pretest-posttest design to test the valence hypothesis and explore the effects of a video demonstration in Phase 2. Our hypothesis is that the observation of a video demonstrating the correct solution of what they have just experienced can result in an improved judgement of stock and flow patterns in subsequent novel problems. Specifically, participants are expected to perform better in a novel SF problem after observing a video that demonstrates that it is possible to control the stock even when the valence of the user-controlled flow opposes the direction of the correct response.

Method

Participants

Participants were recruited through Amazon Mechanical Turk (MTurk) to complete a "Decision Making & Health" task. They were paid US $1.50 for completing the survey and received a bonus of US $0.50 for each question answered correctly up to a maximum payment of US $2.50. Of the 403 recruited participants, two failed to complete the study and two failed to watch the video in full, leaving a total of 399 for analysis. Of these participants ($M_{\text{age}} = 34.25$, $SD_{\text{age}} = 10.14$), 58.40% identified as male, 41.35% identified as female, and 0.25% identified as intersex. Furthermore, 94.49% reported never having suffered from diabetes (not including prediabetes) and 15.04% reported never having helped care for someone with diabetes. Participants were randomly assigned to one of four groups (described below) as follows: 98 participants to Group A, 101 participants to Group B; 101 participants to Group C; and 99 participants to Group D. We did not record the time spent on each part of the experiment, but collected the total time for each participant recorded from Amazon MTurk. The average time was 14 minutes and the median was 12 minutes.

Experimental design

This study was composed of two phases: Phase 1 and Phase 2, divided by the presentation of a video. We used a 2 (stock behaviour: increasing or decreasing) x 2 (flow decision: matching, opposing the correct solution) full factorial design in each phase, that is, before and after showing the video. The first factor was the direction of the accumulation (i.e., stock behaviour), which increased or decreased over time. The *increasing* stock matched the hypoglycemia (low starting blood glucose) scenario, and the *decreasing* stock matched the hyperglycemia (high starting blood glucose level) scenario. The second factor was the valence of the user-controlled flow and whether it matched or opposed the correct response needed to control the blood glucose at a target level at the end of the 100-minute period. The user-controlled flow valence either matched or opposed the direction of the correct response, whereas the other flow was fixed at a constant level.

This 2x2 factorial design yielded four scenarios. Scenario 1 involved an increasing stock (i.e., blood glucose) in which the user-controlled flow matches the increasing direction of the correct response and is "good" (increasing insulin to control the blood glucose); Scenario 2 involved an increasing stock in which the user-controlled flow opposes the increasing direction of the correct response and is "good" (decreasing the consumption of soda to control the blood glucose); Scenario 3 involved a decreasing stock in which the user-controlled flow matches the decreasing direction of the correct response and is "bad" (decreasing insulin to control the blood glucose); and Scenario 4 involved a decreasing stock in which the user-controlled flow opposes the decreasing direction of the correct response and is "bad" (increasing the consumption of soda to control the blood glucose).

	Phase 1		Video	Phase 2	
Group A	Scenario 1		Scenario 1 video	Scenario 2	
	Factor 1:	increasing stock		Factor 1:	increasing stock
	Factor 2:	user-controlled outflow matches the correct solution (outflow should increase)		Factor 2:	user-controlled inflow opposes the correct solution (inflow should decrease)
Group B	Scenario 2		Scenario 2 video	Scenario 1	
	Factor 1:	increasing stock		Factor 1:	increasing stock
	Factor 2:	user-controlled inflow opposes the correct solution (inflow should decrease)		Factor 2:	user-controlled outflow matches the correct solution (outflow should increase)
Group C	Scenario 3		Scenario 3 video	Scenario 4	
	Factor 1:	decreasing stock		Factor 1:	decreasing stock
	Factor 2:	user-controlled outflow matches the correct solution (outflow should decrease)		Factor 2:	user-controlled inflow opposes the correct solution (inflow should increase)
Group D	Scenario 4		Scenario 4 video	Scenario 3	
	Factor 1:	decreasing stock		Factor 1:	decreasing stock
	Factor 2:	user-controlled inflow opposes the correct solution (inflow should increase)		Factor 2:	user-controlled outflow matches the correct solution (outflow should decrease)

Table 1. Experimental groups exposed to one of the four experimental conditions in Phase 1, followed by a video showing a demonstration of a correct solution to the Phase 1 scenario, and then by exposure to a different scenario of experimental conditions in Phase 2.

Appendix 1 shows the four scenarios that were used in the study. Participants were shown two graphs: a graph of the stock trend (i.e., blood glucose) over the course of a 100-minute period, and a graph of the corresponding inflow and outflow (insulin and sugar consumption) trends; one of the flows was fixed at the same level over the 100-minute time period, and the trend of the other flow was shown up to minute 50. Participants were asked to decide how to stabilize the blood glucose level at minute 100 as shown in the stock graph using a sliding bar to decide on the level of the incomplete flow (insulin or sugar consumption).

Experimental groups

Four experimental groups were designed so that participants would receive one of the four Phase 1 scenarios, followed by a video illustrating the dynamics of the solution to the question that they had just been asked to solve in Phase 1. They were then asked to deal with a different scenario in Phase 2. These groups are illustrated in Table 1.

Participants exposed to Scenario 1 in Phase 1 were asked to solve Scenario 2 in Phase 2; those exposed to Scenario 2 in Phase 1 were exposed to Scenario 1 in Phase 2; those exposed to Scenario 3 in Phase 1 were exposed to Scenario 4 in Phase 2, and those exposed to Scenario 4 in Phase 1 were exposed to Scenario 3 in Phase 2. These shifts from Phase 1 to Phase 2 were designed so that we could observe the effect of the video demonstration on the valence effect: the ability of participants to handle scenarios in which the user-controlled flow matches and opposes the direction of the correct response, respectively. They should also enable us to test the valence effect with respect to both increasing and decreasing stock scenarios.

Procedure and video demonstration

After providing informed consent, participants were randomly assigned to one of the four scenarios. Participants were asked to imagine a diabetic person who was trying to control blood glucose (see Appendix 1 for exact instructions). Participants were presented with the Phase 1 scenario and were asked to indicate their flow decision by manipulating a slider that ranged from 0 mg/dL per 10 minutes to 200 mg/dL per 10 minutes. Next, they were shown a video that demonstrated how the problem stated in Phase 1 was correctly solved dynamically. A five-to-six-minute narrated training video was created using the simulated water tank used in Gonzalez & Dutt's (2007) dynamic stock and flows (DSF) task (see Figure 1 for a screenshot of the simulation used in the video). The DSF program displayed a two-dimensional water tank with two inflow pipes (representing the user inflow and the environmental inflow) and two outflow pipes (representing the user outflow and the environmental outflow). The amount associated with each flow was shown numerically on the screen, as was the current amount in the tank. A red horizontal line drawn across the tank represented the target level. At the bottom of the screen, there was a blank field into which the narrator entered the user inflow (outflow) for each time period, where each time period represented 10 minutes. The goal of the task was to adjust the user inflow (outflow) over 10 time periods so that the amount in the tank reached the target level by the end of the tenth time period. The training video presented the same scenario as the problem solved in Phase 1 — but with different values[1]. As

[1] In the DSF program, the unit of time was set by default to "Hour." We blanked out this unit and intended to replace it with "Minute," but we accidentally neglected to do so.

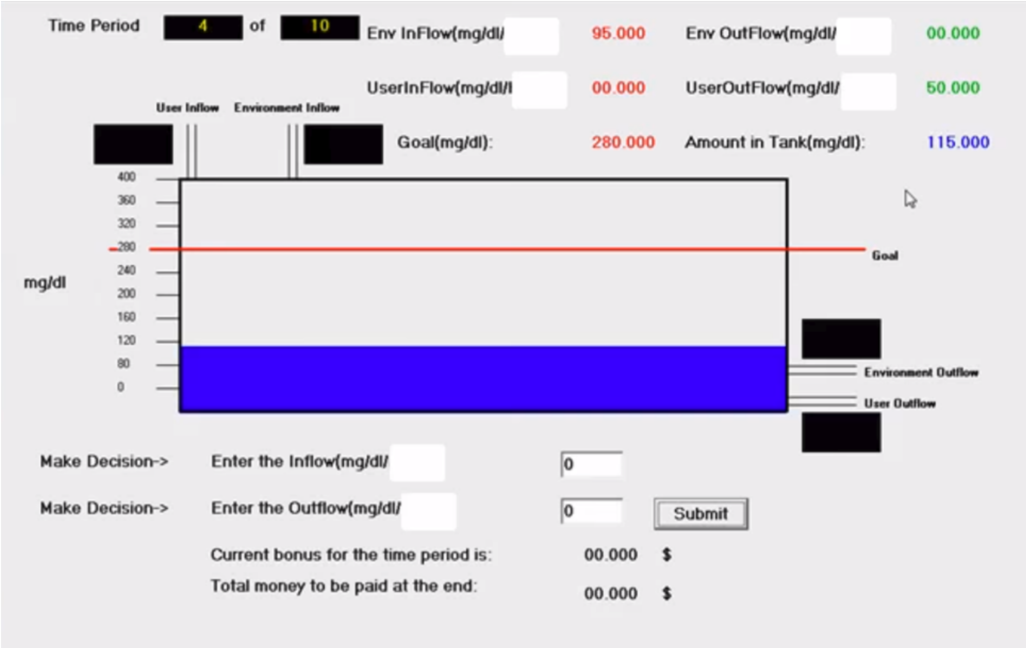


Figure 1. Screenshot of the training video for Group A. The scenario was Scenario 1: hypoglycemia (increasing stock) and the environmental inflow was fixed at 95 mg/dL/minute.

in the graph problems, either the environmental inflow or the environmental outflow was set at a fixed value, whereas the environmental outflow (inflow) was programmed to vary across time periods. The narrator explained each visual component of the DSF task and then acted out the task for the participant, entering values into the user inflow (outflow) box and describing their effect on the level in the tank at each decision point. As the videos progressed, the narrator developed and explained a strategy for reaching the target level (see Appendix 2 for a transcript of the narrator's script). In all four training videos, the narrator attained the goal stock level at the end of the tenth time period.

After watching the training video, the Phase 2 scenario was introduced to participants, and they were asked to make their flow decision using the same slider as in Phase 1. Finally, participants completed a brief demographic questionnaire that assessed the participants' first- and second-hand experience with diabetes (see Appendix 3 for full questionnaire).

Importantly, participants did not receive any direct feedback about the correctness of responses after Phase 1 or after Phase 2. For payment purposes, they were only informed of the total bonus amount they earned at the end of the experiment. Participants were not informed of which of the questions they responded correctly to in which of the two phases.

Results

Performance in both Phase 1 and Phase 2 was measured in terms of the accuracy gap, calculated by subtracting the correct response for each scenario from the participant's response. Therefore, positive values indicated that participants "overshot" the correct response, negative values indicated that participants "undershot" the correct response, values closer to zero indicated higher accuracy, and values that are equal to zero indicate exact correct response. Using this measure, we can quantify the direction and strength of the valence effect. Participants are expected to overshoot the goal when something is "good" and undershoot the goal when something is "bad". Furthermore, when something is "good" to a greater degree, participants are expected to overshoot the goal more than when something is "good" to a lesser degree; similarly, when something is "bad" to a greater degree, participants are expected to undershoot the goal more than when something is "bad" to a lesser degree.

Phase 1 Accuracy

Figure 2 shows the accuracy distributions for each scenario and Table 2 gives the exact measures of central tendency and the proportion of participants who responded correctly in Phase 1.

We found that, regardless of the direction of the stock (increasing or decreasing), participants were more accurate when the valence of the user-controlled flow agrees with the direction of the correct response (Scenarios 1 and 3) than when the user-controlled flow opposes the direction of the correct response (Scenarios 2 and 4). These observations support the correlation heuristic, although large standard deviations were observed in all scenarios. In terms of frequency of correct responses, we found that Scenarios 1 and 3 have higher exact accuracy (21.43% and 16.83% respectively) compared to Scenarios 2 and 4 (8.91% and 9.08% respectively).

In accordance with the valence hypothesis, we found that, on average, participants in Scenarios 1 and 3 overshoot the goal. Thinking that insulin is "good"

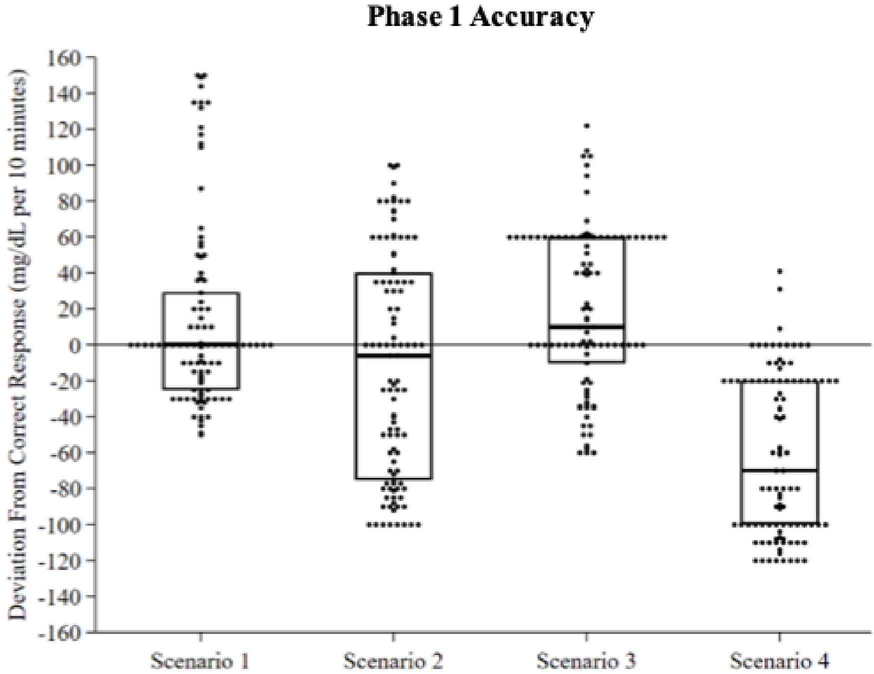


Figure 2. Deviations from the correct response for each of the four scenarios in Phase 1. Each dot in the figure represents a response. Horizontal lines in the box plots represent the first quartile, median, and third quartile. The horizontal line at y = 0 indicates perfect accuracy.

		M	*SD*	*Median*	*Frequency of Correct Response*
Phase 1	**Scenario 1**: increasing stock, user-controlled outflow **matches** the correct solution	13.56	52.30	0.00	21.43%
	Scenario 2: increasing stock, user-controlled inflow **opposes** the correct solution	-11.59	62.77	-6.00	8.91%
	Scenario 3: decreasing stock, user-controlled outflow **matches** the correct solution	19.13	45.00	10.00	16.83%
	Scenario 4: decreasing stock, user-controlled inflow **opposes** the correct solution	-60.61	44.76	-70.00	9.09%

Table 2. Measures of central tendency and the proportion of participants who responded correctly in Phase 1. Values are means, standard deviations, and medians for deviations from the correct response. A value of 0 indicates perfect accuracy. The frequency of correct responses was calculated by dividing the number of participants who gave correct responses in a scenario by the total number of participants in that scenario.

for controlling blood glucose, they used more insulin than necessary to control the increasing blood glucose in Scenario 1 and to control the decreasing blood glucose in Scenario 3. Similarly, on average, participants undershoot the goal in Scenarios 2 and 4. In the belief that soda consumption is “bad” for controlling blood glucose, they used less soda than necessary to control the increasing blood glucose in Scenario 2 and to control the decreasing blood glucose in Scenario 4. Furthermore, as hypothesized by the valence effect, we found that undershooting was greater in Scenario 4 than Scenario 2, that is, the two scenarios in which the user-controlled flow (i.e., soda) opposes the direction of the correct response. People were more reluctant to increase soda consumption (it is “bad”) in order to control the decreasing blood glucose in Scenario 4 than to decrease soda consumption (it is “good”) in order to control for the increasing blood glucose in Scenario 2.

To test these observations, we conducted a 2 (stock behaviour: increasing, decreasing) x 2 (participant-controlled flow: matching, opposing) analysis of variance (ANOVA) on the accuracy scores from Phase 1. This analysis revealed that there was a significant main effect on the flow valence. Participants were more accurate when the valence of the user-controlled flow matched the correct response (Scenarios 1 and 3; $M = 16.39$, $SD = 48.68$) than when the correct flow direction opposed the behaviour of the stock (Scenarios 2 and 4; $M =$ -35.86, $SD = 59.75$), $F(1, 395) = 102.40$, $p < .001$, $\eta_p{}^2 = .21$. Also, stock direction was found have a main effect. Participants were more accurate at increasing stock scenarios (1 and 2; $M = 0.79$,

SD = 59.07) compared to decreasing stock scenarios (3 and 4; M = -20.34, SD = 60.01), $F(1, 395) = 17.57$, $p < .001$, $\eta_p^2 = .04$.

Furthermore, a significant interaction effect emerged between the user-controlled flow and the stock direction, $F(1, 395) = 27.73$, $p < .001$, $\eta_p^2 = .07$. As shown in Figure 3, while there was no difference in the degree of overshooting in scenarios in which the user-controlled flow valence matched the stock direction (Scenarios 1 and 3), there was a significant difference in the amount of undershooting in scenarios in which the valence of the participant-controlled flow opposed the correct solution (Scenarios 2 and 4). These observations were confirmed with simple effects tests using the Bonferroni correction to adjust for multiple comparisons. Participants in Scenario 4 undershot significantly more (M = -60.61, SD = 44.76) than participants in Scenario 2 (M = -11.59, SD = 62.77), $p < .001$. No significant difference in accuracy emerged between participants in Scenario 1 and Scenario 3, $p = .45$.

Phase 2 Accuracy

Figure 4 illustrates the accuracy distributions for each scenario and Table 3 gives the exact measures of central tendency and the proportion of participants who responded correctly in Phase 2. The observed results are very similar to findings in Phase 1. Again we observe that the valence of the user-controlled flow has an effect: accuracy is higher in Scenarios 1 and 3 (21.78% and 28.28%, respectively) than in Scenarios 2 and 4 (23.47% and 9.09%, respectively). We also observed overshooting and undershooting of the goal in Scenarios 1 and 3 and Scenarios 2 and 4, respectively. Again undershooting was greater in Scenario 4 than in Scenario 2.

We conducted a 2 x 2 ANOVA on Phase 2 accuracy scores. As in Phase 1, a significant main effect of flow valence emerged: $F(1, 395) = 112.84$, $p < .001$, $\eta_p^2 = .22$. Specifically, participants performed better in Scenarios 1 and 3 (M = 10.27, SD = 34.75) than in Scenarios 2 and 4 (M = -31.73, SD = 46.37). Also, stock direction had a significant main effect, where participants who were shown an increasing stock (Scenarios 1 and 2; M = -0.41, SD = 42.86) performed better than those shown a decreasing stock (Scenarios 3 and 4) (M = -20.90, SD = 47.04), $F(1, 395) = 25.92$, $p < .001$, $\eta_p^2 = .06$. There was also a significant interaction between the user-controlled flow and the stock direction: $F(1, 395) = 12.17$, $p = .001$, $\eta_p^2 = .03$. Figure 5 illustrates this interaction. Again, there was a significant difference in the degree of undershooting between Scenarios 2 and 4, in which the participant-controlled flow valence opposed the correct solution, although there was no significant difference in the degree of overshooting between Scenarios 1 and 3, in which the participant-controlled flow valence matched the correct solution. To confirm these observations, a test of simple effects was performed using the Bonferroni correction to adjust for multiple comparisons. Participants in Scenario 4 undershot significantly more (M = -48.33, SD = 43.71) than participants in Scenario 2 (M = -14.63, SD = 42.86), $p < .001$. There was no significant difference in the amount of overshooting between participants in Scenario 1 and Scenario 3, $p = .26$.

Effects of video demonstration and demographics

While the separate analyses of Phase 1 and Phase 2 reported above suggest that there is little or no difference between conditions in Phase 1 and Phase 2, this section describes an analysis of the effects for each participant group (see Table 1). We tested whether accuracy of participants in Phase 2 would improve after a video demonstration illustrating the correct response to the scenario experienced in Phase 1. Of particular interest is the potential improvement in Groups A and C, where participants perform Scenarios 2 and 4 in Phase 2, respectively, because they are the scenarios in which the user-controlled flow opposes the correct solution, where accuracy was found to be lower.

We compared average improvements in accuracy between Phase 1 and Phase 2 across conditions. Improvement was calculated by subtracting the absolute value of deviations from the correct response in Phase 2 from the absolute value of deviations from the correct response in Phase 1 for each condition. More positive values indicate improvement, whereas more negative values indicate degradation. On average, across all conditions, participant accuracy improved by 14.80 units (SD = 52.29) from Phase 1 to Phase 2 after the presentation of the video.

Figure 6 presents the average improvement for each of the four groups. We performed a 2 (stock direction in video: increasing, decreasing) x 2 (participant-controlled flow in video: matching, opposing) ANOVA[2] on accuracy improvement. A significant main effect of user-controlled flow valence emerged: $F(1, 395) = 64.18$, $p < .001$, $\eta_p^2 = .14$. Specifically, there was a significant improvement for participants in groups B and D from Phase 1 to Phase 2 (M = 34.14, SD = 47.29), while performance for participants in Groups A and C dropped from Phase 1 to Phase 2: M = -4.63, SD = 49.92. Stock direction did not have a significant main effect: $p = .66$. However, the interaction between flow valence and stock direction was significant: $F(1, 395) = 7.88$, $p = .005$, $\eta_p^2 = .02$. Tests of simple effects with Bonferroni corrections revealed that the accuracy for participants in Group A (solving Scenario 1 in Phase 1 and Scenario 2 in Phase 2) improved (M = 3.34, SD = 50.44), whereas accuracy dropped in Group C (solving Scenario 3 in Phase 1 and Scenario 4 in Phase 2) (M = -12.37, SD = 48.41): $p = .02$. There was no significant difference in improvement between Group B and Group D: $p = .10$. In the presence of

[2] Recall that the stock behavior and flow decision in the training video matched the stock behavior and flow decision that participants saw in Phase 1.

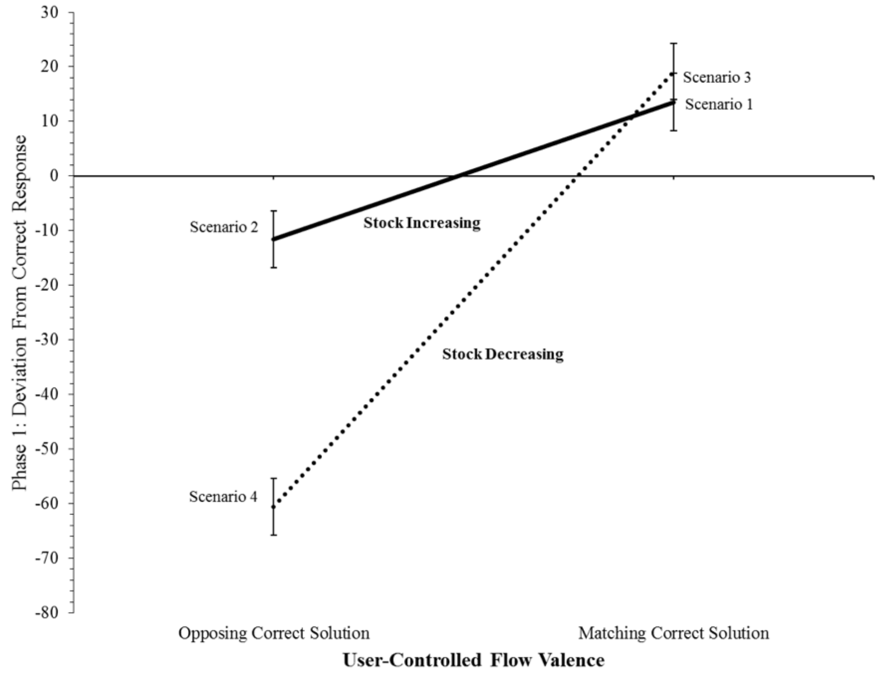


Figure 3. Significant interaction between flow decision (matching, opposing) and stock behaviour (increasing, decreasing) in Phase 1. Deviations were measured in mg/dL per 10 minutes. Error bars indicate standard errors of the mean.

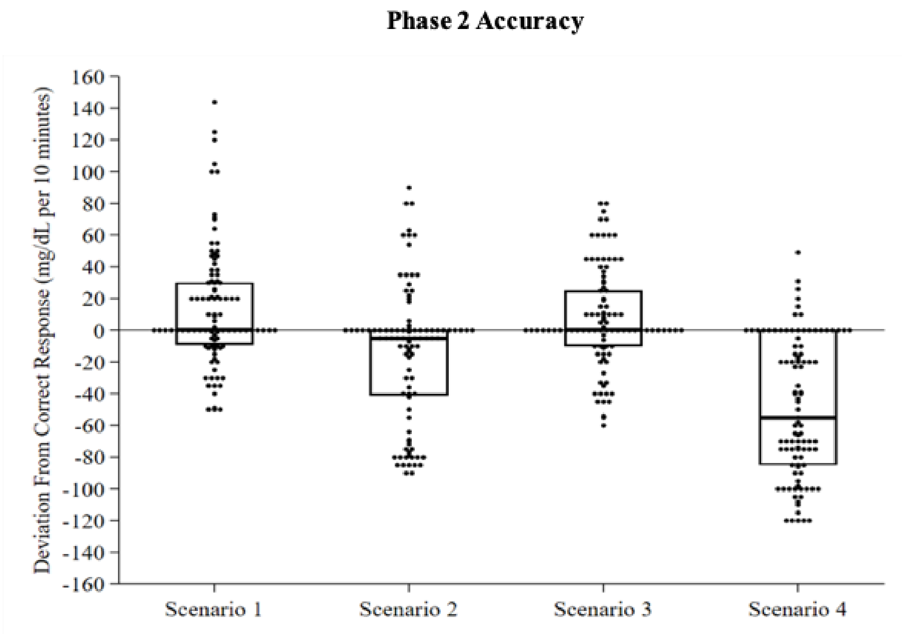


Figure 4. Deviations from the correct response for each of the four scenarios in Phase 2. Each dot in the figure represents a response. Horizontal lines in the box plots represent the first quartile, median, and third quartile. The horizontal line at y = 0 indicates perfect accuracy. (see Figure 2).

a video demonstration there were significant improvements for groups B and D (i.e., in scenarios with inflows matching the correct solution after working on scenarios with inflows opposing the correct solution).

Finally, a demographic analysis regarding factors of possible interest — educational level, being a diabetes sufferer, and diabetes caring experience — showed that none of these factors influenced the accuracy of participants' responses in Phase 1 or Phase 2 or the improvement from Phase 1 to Phase 2. No variables were significant.

Discussion

Overall, we found that the valence of the user-controlled flow and the stock direction had significant main effects, and there was an interaction between valence and stock direction influencing accuracy in Phase 1 and Phase 2. Regardless of the direction of the stock (increasing, decreasing), participants responded more accurately when the valence of the user-controlled flow matched the direction of the correct solution to the problem than when the valence op-

		M	*SD*	*Median*	*Frequency of Correct Response*
Phase 2	**Scenario 1**: increasing stock, user-controlled outflow **matches** the correct solution	13.39	37.76	0.00	21.78%
	Scenario 2: increasing stock, user-controlled inflow **opposes** the correct solution	-14.63	42.86	-5.00	23.47%
	Scenario 3: decreasing stock, user-controlled outflow **matches** the correct solution	7.09	31.25	0.00	28.28%
	Scenario 4: decreasing stock, user-controlled inflow **opposes** the correct solution	-48.33	43.71	-55.00	9.09%

Table 3. Measures of central tendency and the proportion of participants who responded correctly in Phase 1. Values are means, standard deviations, and medians for deviations from the correct response. A value of 0 indicates perfect accuracy. The frequency of correct responses was calculated by dividing the number of participants who gave the correct responses in a scenario by the total number of participants in the corresponding group (see Table 2).

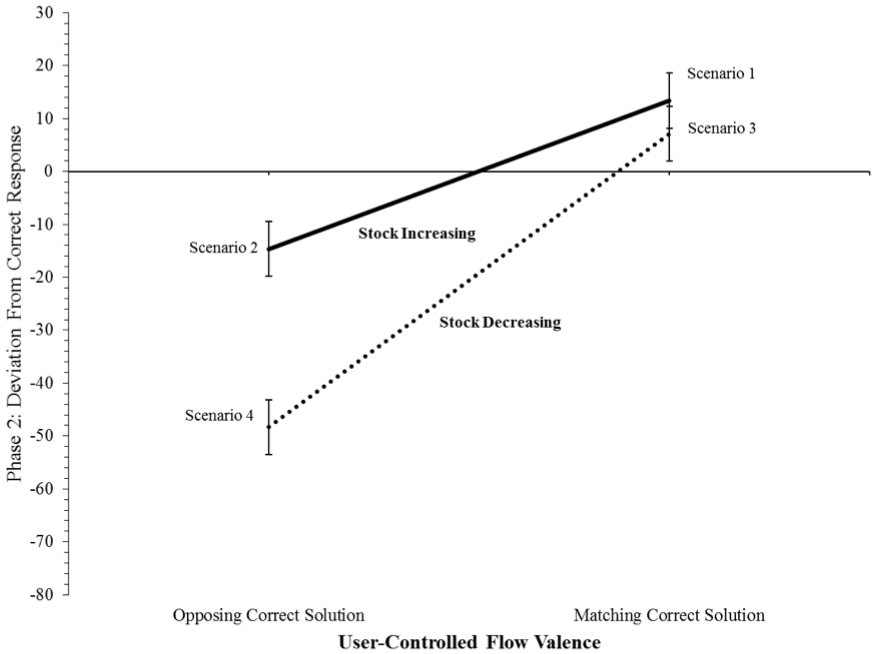


Figure 5. Significant interaction between flow decision (matching, opposing) and stock behaviour (increasing, decreasing) in Phase 2. Deviations were measured in mg/dL per 10 minutes. Error bars indicate standard errors of the mean.

posed the direction of the correct response. This finding is consistent with research that demonstrates that people more easily learn positive relationships among variables (i.e., variables moving in the same direction) than negative relationships (i.e., variables moving in opposing directions) (Brehmer, 1980). It also provides strong support for the correlation heuristic (Cronin, et al., 2009).

These results also corroborate and expand the results reported by Newell et al. (2016) regarding the valence effect. First, we replicated Newell et al.'s findings with respect to the valence effect on increasing stocks. When the user-controlled flow valence matched the direction of the correct solution (Scenario 1, i.e., increase of insulin to control the blood glucose accumulation), performance was better than when the user-controlled flow valence opposed the direction of the correct solution (Scenario 2, i.e., decrease soda consumption to control the blood glucose accumulation). Second, we also expanded this result to decreasing stocks. We found that when the valence matched the direction of the correct solution (Scenario 3, i.e., decrease insulin in order to control the blood glucose accumulation), performance was better than in Scenario 4 (i.e., increase soda consumption in order to control the blood glucose accumulation). As suggested by Newell et al. (2016), performance was found to be good not as a result of the actual understanding of stock-flow relationships but because the valence of the user-controlled flow matched the direction of the correct solution and the correlation heuristic reinforced this relationship. Consistently, and regardless of the direction of the stock, participants responded worse to situations where the valence of the user-controlled flow opposed the direction of the correct solution, that is, only a few participants were able to overcome the correlation heuristic.

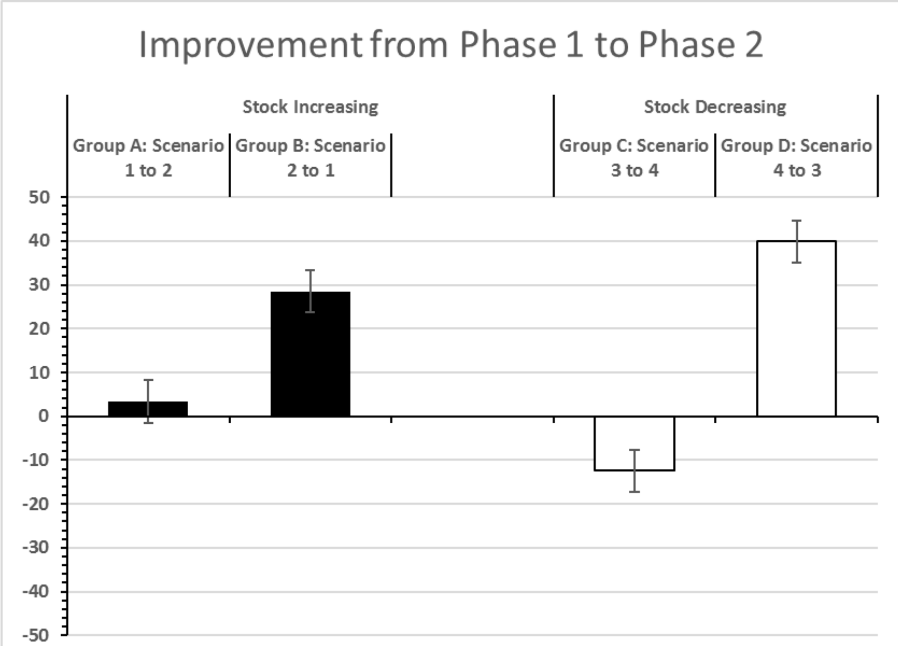


Figure 6. Mean improvement in accuracy between Phase 1 and Phase 2 for the four experimental groups. Error bars represent standard errors of the mean.

Furthermore, the stock direction also had a main effect. In both Phase 1 and Phase 2, participants responded more accurately to problems involving increasing stocks (i.e., Scenarios 1 and 2) than to problems involving decreasing stocks (i.e., Scenarios 3 and 4), a result that supports past findings that increasing relationships are easier to learn than decreasing ones (Gonzalez & Dutt, 2007). However, a significant interaction explains the relationships between the valence of the problem and the direction of the stock. There is significant amount of overshooting of the goal in Scenarios 1 and 3 and a significant amount of undershooting the goal in Scenarios 2 and 4. According to the valence effect, thinking that the use of insulin is "good" for controlling blood glucose resulted in the use of more insulin than needed, whereas participants considering that soda ingestion is "bad" for controlling blood glucose ended up using less soda than needed to achieve the target blood glucose level. Interestingly, and as predicted by the valence hypothesis, we found significantly larger undershooting in Scenario 4 than in Scenario 2. This is explained by the interaction between the direction of the stock and the opposition of the valence of the user-controlled flow. In Scenario 2, the blood glucose level increases over time. This is "bad", and a decrease in the soda consumption would appear to be a "good" option for counteracting this trend. Thus, there is undershooting in Scenario 2, but it is less severe than in Scenario 4. In Scenario 4, the blood glucose level decreases over time, this is "good" and an increase in the soda consumption would appear to be a "bad" option for counteracting this trend, increasing the undershooting of the goal. In actual fact, the correct solution is to decrease and increase soda consumption in Scenarios 2 and 4, respectively.

Regarding the improvement in accuracy from Phase 1 to Phase 2, while the observed improvements cannot be attributed exclusively to the observation of the video demonstration, it is interesting to discuss the different effects and how they relate to the valence effect. We observed a substantial, significant and similar improvement for participants in groups B and D, a slight improvement for participants in Group A; and a deterioration of performance in Group C. One explanation for the improvement observed in Groups B and D is that the participants switched from Scenarios in which solutions *opposed* the correlation heuristic in Phase 1 to Scenarios whose solutions *matched* the correlation heuristic in Phase 2. In general, participants performed more accurately in Phase 2 than in Phase 1, but this improvement could be due to the use of the correlation heuristic rather than to the video demonstration. Therefore, the improvement in performance from Phases 1 to 2 may be fortuitous as in the case reported by Newell et al. (2016). As the correlation heuristic is a robust tendency, it is unlikely, in view of the results for Groups A and B, that the observed improvement can be attributed to the video demonstration alone.

Group A participants performed more accurately in Scenario 2 of Phase 2 than they did in Scenario 1 of Phase 1. This is interesting as it is consistent with the idea that the video demonstration may have influenced the participants to accept the decrease in soda consumption (a "good" action) to stabilize an increasing blood glucose level in Scenario 2. However, the video demonstration did not clearly influence accuracy in Scenario 4 after the demonstration of Scenario 3. Although the undershooting in Scenario 4 of Phase 2 decreased compared to Phase 1, the drop is insufficient to suggest that participants considered the increase in soda consumption as a "good" option for stabilizing a decreasing blood glucose level.

Implications, Limitations and Future Work

A main implication of these results is that one needs to carefully select how to frame the context of a stock and flow problem. To encourage correct responses, one needs to: (1) frame a problem so that the correct response matches the correlation heuristic; (2) ensure that the valence of the scenario matches the valence of the correct response if the correct response cannot match the correlation heuristic; and (3) avoid situations in which the correct response matches neither the correlation heuristic nor the valence of the scenario, because this will result in extremely poor performance. A demonstration of the first strategy is discussed in Dutt and Gonzalez (2013). This "information presentation" strategy proposed to take advantage of the human reliance on the correlation heuristic to encourage people to pay eco-taxes. Participants judged that larger eco-tax increases would cause proportionally greater reductions in CO_2 emissions yet preferred smaller tax increases because of their lesser cost. This finding suggests that it would be beneficial for eco-tax policy makers to present information in terms of eco-tax increases such that smaller than current eco-tax increases (which are more attractive and are likely to be chosen by people) cause greater CO_2 emissions reductions. In future research, this practical conclusion and the effects of the scenario valence matching or opposing the direction of the correct response need to be tested in naturalistic cases of blood glucose control and other global problems.

Also, it is important in future research to address the limitations of the design of this study to test the effectiveness of the video demonstration and possible interventions to improve performance in SF tasks. As discussed above, the improvement of the performance from Phase 1 to Phase 2 that we observed in some of the groups (B and D) may have been accidental, as in the case reported in Newell et al. (2016), especially considering the fact that there was little improvement in Group A and a decline in performance for Group C. In addition, the observed improvements may not be exclusively due to the video demonstration. There could be a practice effect, due to the exposure to a similar (albeit not identical) problem in Phase 1. Future research should focus on experimentally manipulating decision aids or interventions in the form of video demonstrations or dynamic simulations, including a control condition, namely, the absence of such intervention. In this manner, we could test the effect of a video demonstration as an intervention for improving accuracy in SF failure.

Acknowledgements: This work has been funded by the National Science Foundation, Decision Risk and Management Science, Award Number 1530479. We thank Research Assistant, Nalyn Sriwattanakomen, who helped with initial preparation of materials and data collection. *Note*: an older version of this work appeared as an abstract in the proceedings of the 36th International Conference of the System Dynamics Society, Reykjavik, Iceland, August 6-10, 2018.

Declaration of conflicting interests: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be constructed as a potential conflict of interest.

Author contributions: Gonzalez contributed to the development ideas, experimental design, materials and data collection, data analyses, and writing. Sanchez-Segura, Dugarte-Peña, and Medina-Dominguez contributed to data analyses, interpretations of results, and writing.

Handling editor: Andreas Fischer

Citation: Gonzalez, C., Sanchez-Segura, M.-I., Dugarte-Peña, G.-L., & Medina-Dominguez, F. (2018). Valence Matters in Judgments of Stock Accumulation in Blood Glucose Control and Other Global Problems. *Journal of Dynamic Decision Making, 4*, 3. doi:10.11588/jddm.2018.1.49607

Received: 19 Jul 2018
Accepted: 18 Dec 2018
Published: 26 Dec 2018

References

Abdel-Hamid, T., Ankel, F., Battle-Fisher, M., Gibson, B., Gonzalez-Parra, G., Jalali, M., ... Murphy, P. (2014). Public and health professionals' misconceptions about the dynamics of body weight gain/loss. *System Dynamics Review*, *30*(1–2), 58–74. doi:10.1002/sdr.1517

Booth Sweeney, L. & Sterman, J.D. (2000). Bathtub dynamics: initial results of a systems thinking inventory. *System Dynamics Review*, *16*, 249–286. doi:10.1002/sdr.198

Brehmer, B. (1980). In one word: Not from experience. *Acta Psychologica*, *45*(1-3), 223–241. doi:10.1016/0001-6918(80)90034-7

Brunstein, A., Gonzalez, C., & Kanter, S. (2010). Effects of domain experience in the stock–flow failure. *System Dynamics Review*, *26*(4), 347–354. doi:10.1002/sdr.448

Cronin, M., Gonzalez, C., & Sterman, J.D. (2009). Why don't well-educated adults understand accumulation? A challenge to researchers, educators and citizens. *Organizational Behaviour and Human Decision Processes*, *108*, 116–130. doi:10.1016/j.obhdp.2008.03.003

Diehl, E., & Sterman, J. D. (1995). Effects of feedback complexity on dynamic decision making. *Organizational Behavior and Human Decision Processes*, *62*(2), 198 –215. doi:10.1006/obhd.1995.1043

Dutt, V., & Gonzalez, C. (2012a). Human control of climate change. *Climatic Change*, *111*(3–4), 497–518. doi:10.1007/s10584-011-0202-x

Dutt, V., & Gonzalez, C. (2012b). Decisions from experience reduces misconceptions about climate change. *Journal of Environmental Psychology*, *32*(1), 19–29. doi:10.1016/j.jenvp.2011.10.003

Dutt, V., & Gonzalez, C. (2013). Enabling eco-friendly choices by relying on the proportional-thinking heuristic. *Sustainability*, *5*(1), 357–371. doi:10.3390/su5010357

Fischer, H., Degen, C., & Funke, J. (2015). Improving stock-flow reasoning with verbal formats. *Simulation & Gaming*, *46*(3–4), 1–15. doi:10.1177/1046878114565058

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*, 496–512. doi:10.1111/cogs.12239

Frensch, P., & Funke, J. (Eds.). (1995). *Complex problem solving: The European perspective*. Hillsdale, NJ: Lawrence Erlbaum Associates. Retrieved from http://works.bepress.com/joachim_funke/13/

Gonzalez, C., & Dutt, V. (2007). Learning to control a dynamic task: A system dynamics cognitive model of the slope effect. *In Proceedings of the ICCM Eighth International Conference on Cognitive Modeling*, Ann Arbor, MI.

Gonzalez, C., & Wong, H. (2012). Understanding stocks and flows through analogy. *System Dynamics Review*, *28*(1), 3–27. doi:10.1002/sdr.470

Moxnes, E., &Saysel, A.K. (2009). Misperceptions of global climate change: Information policies. *Climatic Change*, *93*(1–2), 15–37. doi:10.1007/s10584-008-9465-2

Newell, BR., Kary, A., Moore, C., & Gonzalez, C. (2016). Managing the budget: Stock-flow reasoning and the CO_2 accumulation problem. *Topics in Cognitive Science*, *8*(1), 138–159. doi:10.1111/tops.12176

Sterman, J.D. (2008). Risk communication on climate change: Mental models and mass balance. *Science*, *322*, 532–533. doi:10.1126/science.1162574

Sterman, J.D., & Booth Sweeney L. (2002). Cloudy skies: Assessing public understanding of global warming. *System Dynamics Review*, *18*(2), 207–240. doi:10.1002/sdr.242

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*(3–4), 213–238. doi:10.1007/s10584-006-9107-5

Sterman, J.D., Fiddaman, T., Franck, T., Jones, A., McCauley, S., Rice, P., . . . Siegel, L. (2012). Management flight simulators to support climate negotiations. *Environmental Modeling and Software*, *44*, 122–135. doi:10.1016/j.envsoft.2012.06.004

Sternberg, R. J. (1995). Expertise in complex problem solving: A comparison of alternative conceptions. In P. A. Frensch & J. Funke (Eds.), *Complex problem solving: The European perspective* (pp. 295–321). Hillsdale, NJ: Lawrence Erlbaum Associates.

Strohhecker, J., Größler, A. (2015). Performance in tangible and in cognitive stock-flow tasks: Closer than expected. *Simulation & Gaming*, *46*(3-4), 230–254. doi:10.1177/1046878115577160

Appendix 1

This appendix describes the four scenarios used in the experiment. All graphs illustrate information on a stock accumulation or a flow. In all cases, the independent variable (shown on the x-axis) is the time (measured in minutes), and the dependent variable is either the stock accumulation value (measured in mg/dL) or the inflow or outflow value (measured in mg/dL/min). The stock accumulation case (hypo- or hyperglycemia) and the explanation of the related fixed and estimated flows (soda consumption – glucose input – or insulin action – glucose output–) are described for each scenario.

Graph Problems

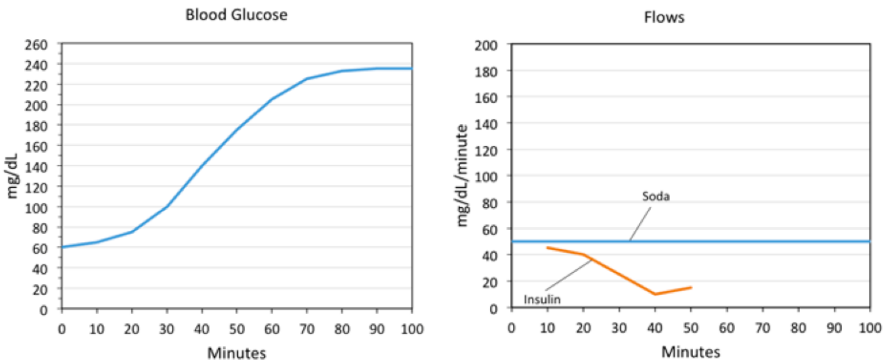
Scenario 1: Hypoglycemia (Increasing Stock) x User-Controlled Outflow

Below are two graphs representing a hypothetical scenario of blood glucose control. The graph on the left shows your **blood glucose levels over a 100-minute period**. The graph on your right shows the rate of your **insulin production, which reduces your blood glucose**, and the rate of your **soda consumption, which increases your blood glucose**.

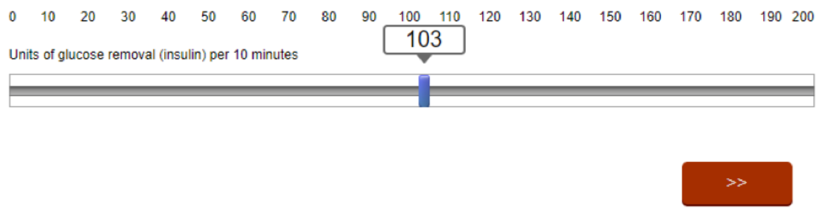
Suppose that you have a variety of diabetes that makes your body produce insufficient insulin.

While exercising, you begin to feel very faint. You measure your blood glucose at 0 minutes and find that it is extremely low at 60 mg/dL, so you ingest a bottle of sugary soda that will continuously put 50 mg/dL of glucose per minute into your bloodstream for the next 100 minutes. However, you soon realize that if you do not start burning off glucose, your blood glucose will become dangerously high.

To remedy this problem, you begin injecting insulin at 10 minutes that removes glucose from your bloodstream. The rate at which you inject insulin is shown for only the first 50 minutes.



If your goal is to **stabilize your blood glucose at 235 mg/dL by 100 minutes** as shown in the left graph, what does the rate of glucose removal (by insulin injection) need to be at 100 minutes?

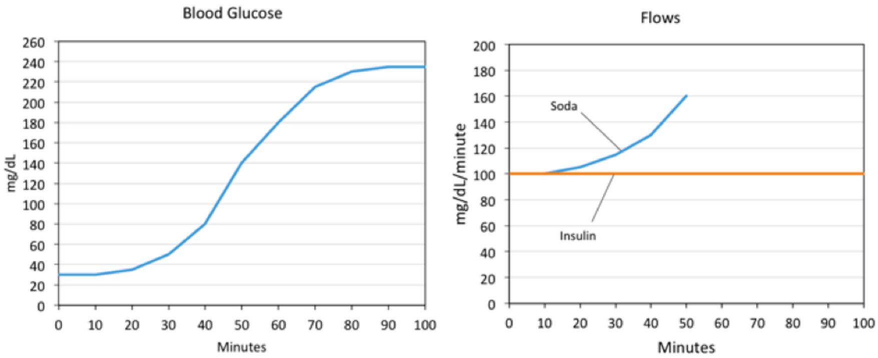


Scenario 2: Hypoglycemia (Increasing Stock) x User-Controlled Inflow

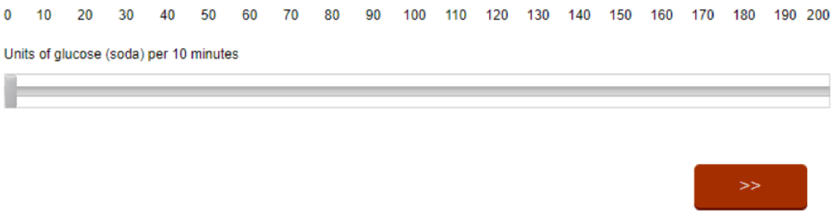
Below are two graphs representing a hypothetical scenario of blood glucose control. The graph on the left shows your **blood glucose levels over a 100-minute period**. The graph on your right shows the rate of your **insulin production, which reduces your blood glucose**, and the rate of your **soda consumption, which increases your blood glucose**.

Suppose that you have a variety of diabetes that prevents your body from producing insulin.

While exercising, you begin to feel very faint. Before you started, you had connected yourself to an insulin pump that will continuously remove 100 mg/dL of glucose per minute from your bloodstream for the next 100 minutes. Now, at 0 minutes, you find that your blood glucose is extremely low at 30 mg/dL.



To remedy this problem, you begin drinking a sugary soda at 10 minutes that puts glucose into your bloodstream. The rate at which you intake glucose through soda is shown for only the first 50 minutes.
If your goal is to **stabilize your blood glucose at 235 mg/dL** by 100 minutes as shown in the left graph, what does the rate of glucose intake (by soda) need to be at 100 minutes?



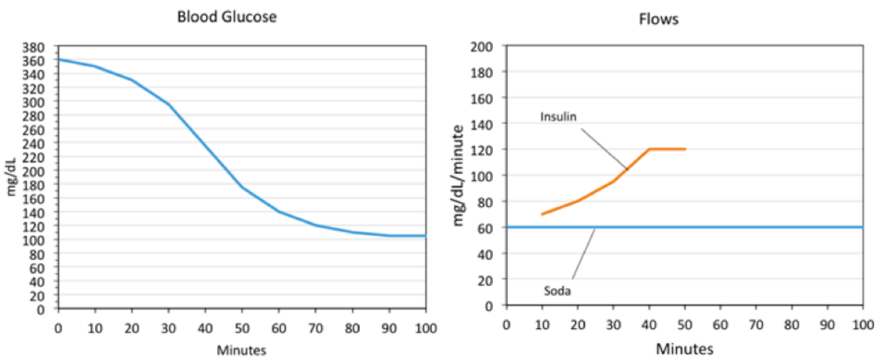
Scenario 3: Hyperglycemia (Decreasing Stock) x User-Controlled Outflow

Below are two graphs representing a hypothetical scenario of blood glucose control. The graph on the left shows your **blood glucose levels over a 100-minute period**. The graph on your right shows the rate of your **insulin production, which reduces your blood glucose**, and the rate of your **soda consumption, which increases your blood glucose**.

Suppose that you have a variety of diabetes that makes your body produce insufficient insulin.

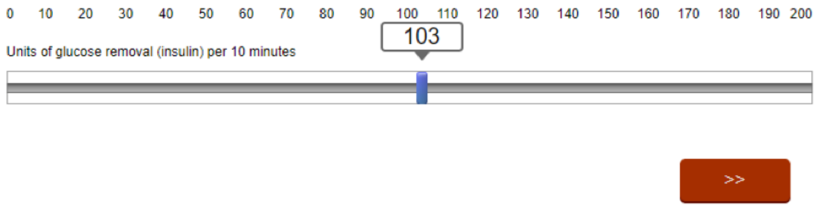
You are drinking a bottle of sugary soda that will continuously put 60 mg/dL of glucose per minute into your bloodstream for the next 100 minutes. You measure your blood glucose at 0 minutes and find that it is extremely high at 360 mg/dL. You soon realize that if you do not start burning off glucose, you will need to be hospitalized.

To remedy this problem, you begin injecting insulin at 10 minutes that removes glucose from your bloodstream. The rate at which you inject insulin is shown for only the first 50 minutes.



If your goal is to **stabilize your blood glucose at 105 mg/dL** by 100 minutes as shown in the left graph, what does the rate of glucose removal (by insulin injection) need to be at 100 minutes?
Scenario 4: Hyperglycemia (Decreasing Stock) x User-Controlled Inflow

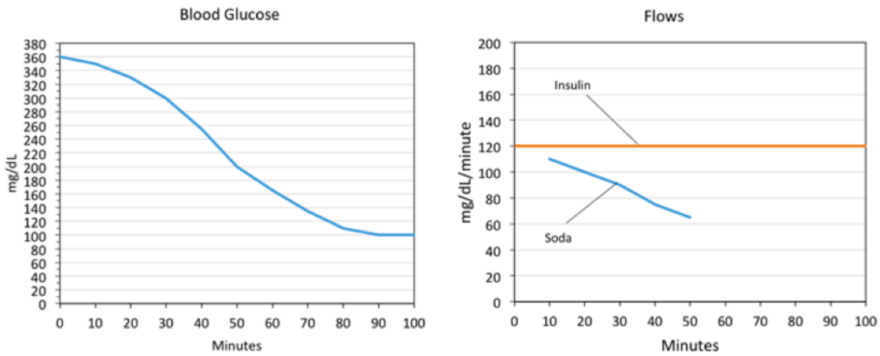
Below are two graphs representing a hypothetical scenario of blood glucose control. The graph on the left shows your **blood glucose levels over a 100-minute period**. The graph on your right shows the rate of your **insulin production, which reduces your blood glucose**, and the rate of your **soda consumption, which increases your blood glucose**.



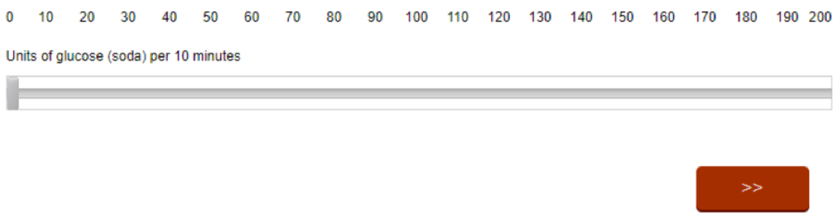
Suppose that you have a variety of diabetes that prevents your body from producing insulin.

You measure your blood glucose at 0 minutes and find that it is extremely high at 360 mg/dL, so you put yourself on an insulin pump that will continuously remove 120 mg/dL of glucose per minute from your bloodstream for the next 100 minutes. However, you soon realize that if you do not start ingesting sugar, your blood glucose will end up dangerously low.

To remedy this problem, you start drinking a sugary soda at 10 minutes that puts glucose into your bloodstream. The amount of glucose that you intake through soda is shown for only the first 50 minutes.



If your goal is **stabilize your blood glucose at 100 mg/dL** by 100 minutes as shown in the left graph, what does the amount[3] of glucose intake (by soda) need to be at 100 minutes?



[3]This term is misleading; it should be "rate," not "amount." This confusion between "amount" and "rate" might account for some of the observed incorrect responses.

Appendix 2

Narrator Script for Video in Scenario 1 (Increasing Stock x User-Controlled Outflow)

Here we have a tank that symbolizes the human body. The blue liquid in this tank stands for the amount of glucose in the bloodstream. We are looking at the bloodstream of a diabetic person who is hypoglycemic. Right now, the blood glucose is very low at 55 mg/dl, and we want to increase it. The red line stands for the target level, which is 280 mg/dl. This goal is definitely high, but for someone who is about to exercise and use up blood glucose in the process, 280 mg/dl is a reasonable level.

We will be controlling the outflow of glucose into the blood by entering numbers into this field here. We're going to be doing this over a span of 100 minutes that will be broken down into 10 periods of 10 minutes each. At each of the 10 time periods, we can choose how much glucose to remove from the body. The glucose will leave through this tube here.

However, the human body has its own checks and balances. Hormones can release extra glucose into the bloodstream through the tube marked "Environment Inflow," and hormones can also absorb glucose through the tube marked "Environment Outflow." We don't know the rate at which the body will receive or absorb glucose, but we'll try at each step to get the glucose level closer to that red line—the target glucose level.

Let's start out by removing, uh, 105 units of glucose from the body. Some glucose is entering the body through this tube, and then we see that the 105 units of glucose we entered is leaving the body here. It looks like 95 units of glucose entered the body and there are now 45 mg/dl. We removed more glucose than the body received, so the blood glucose level decreased over this first time period! That's not what we want.

So 10 minutes have passed. It's time for us to make a decision about how much glucose to remove during the next 10 minutes. Let's try removing 95 units. Some glucose is entering the body through this tube, and then we see that the 95 units of glucose we entered is leaving the body here. It looks like another 95 units of glucose entered the body and there are now 45 mg/dl. We removed as much glucose as the body received, so the blood glucose level remained the same!

All right, another 10 minutes have passed. Let's make a decision about how much glucose to remove for the next time period. Let's try removing 70 units. Glucose is entering the body through here, and then the 70 units are leaving from here. So, it appears that another 95 units of glucose entered the body and there are now 70 mg/dl. We removed less glucose than the body received, so the blood glucose level increased! We're getting a little closer to the target level.

Time to make another decision about how much glucose to remove during the next time period. Let's try removing even less than before: 50 units. Glucose enters the body here, and then the 50 units leave. Again, another 95 units of glucose entered, and there are now 115 mg/dl. We removed less glucose than the body received, so the level is continuing to increase. We're getting even closer!

Time to make another decision about how much glucose to remove during the next time period. Let's try removing less than before: 20 units this time. Glucose is entering; 20 units leave. Again, another 95 units of glucose entered, and there are now 190 mg/dl. Because we removed less glucose than the body received, the level is increasing. We're halfway through the 100-minute time span now!

This time, let's try removing more than last time: 45 units. Glucose is entering; 45 units leave. Again, 95 units of glucose entered, and there are now 240 mg/dl! Because we still removed less glucose than the body received, the level increased. We're much closer to the target now.

This time, let's remove more glucose: 70 units. Glucose enters; 70 units leave. Yes, 95 units of glucose entered, and there are now 265 mg/dl. The same principle applies: if the glucose removed is less than the amount received, the overall level in the body increases.

All right, let's try removing even more: 85 units. Glucose enters, our 85 units leave, and now there are 275 mg/dl. That's very, very close to the target amount.

We should continue to increase the amount of glucose we're removing. Otherwise, we'll undershoot the goal. So let's remove 90 units. 95 units enter—the blood glucose is too high—but then our 95 units leave, and we're at exactly 280 mg/dl. Perfect! Now all we need to do is maintain it at that level for the final time period.

So we've seen that the body receives 95 units during each time period. If we want to keep the blood glucose at the same level, we have to remove exactly as much as the body receives; the two amounts will cancel out. So let's remove 95 units. Yes, 95 enter, 95 leave, and yes! We're at exactly 280 mg/dl at the end of the 100-minute span. Success!

Appendix 3

Demographic Questionnaire

1. What is your sex? [Choose one.]
   a. Male [58.40%]
   b. Female [41.35%]
   c. Intersex [0.25%]
2. What is your age? [Free response.] (M = 34.25, SD = 10.14)
3. What is the highest level of education that you have completed? [Choose one.]
   a. Some high school [0.75%]
   b. High school [13.28%]
   c. Some college [26.32%]
   d. Associate's degree [13.53%]
   e. Bachelor's degree [38.60%]
   f. Master's degree [5.76%]
   g. Professional or doctoral degree [1.75%]
4. Do you have or have you ever had any form of diabetes (NOT including prediabetes)? [Choose one.]
   a. Yes [3.01%]
   b. No [94.49%]
   c. I don't know [2.51%]
5. What kind of diabetes do you have or have you ever had? (Please check all that apply.) [This question was shown only if the participant responded "Yes" to Question 4.]
   a. Type 1 [8.31% of those who said "Yes"]
   b. Type 2 [75.08%]
   c. Gestational [24.92%]
6. When were you diagnosed with diabetes? [Choose one. This question was shown only if the participant responded "Yes" to Question 4.]
   a. Less than a year ago [16.67% of those who said "Yes"]
   b. 1 to 3 years ago [16.67%]
   c. 4 to 6 years ago [25.00%]
   d. 7 or more years ago [41.67%]
7. Are you currently receiving treatment or have you ever received treatment for your diabetes? Treatments include insulin shots, insulin pumps, and medications. [Choose one. This question was shown only if the participant responded "Yes" to Question 4.]
   a. Yes [83.33% of those who said "Yes" to Question 4]
   b. No, but I have made lifestyle changes [16.67%]
   c. No, and I have not made lifestyle changes [0.00%]
8. If you currently have diabetes, how well-controlled would you say it is? [Choose one. This question was shown only if the participant responded "Yes" to Question 4.]
   a. Extremely well [0.00% of those who said "Yes" to Question 4]
   b. Very well [41.67%]
   c. Moderately well [0.00%]
   d. Very poorly [16.67%]
   e. Extremely poorly [25.00%]
   f. I do not currently have diabetes [16.67%]

9. Does anyone who is close to you (such as a relative, a spouse, or a close friend) have diabetes? [Choose one.]
   a. Yes [37.59%]
   b. No [62.41%]
10. Have you ever helped this person(s) manage their diabetes? [Choose one. This question was shown only if the participant responded "Yes" to Question 9.]
    a. Yes [28.67% of those who said "Yes" to Question 9]
    b. No [71.33%]
11. Have you ever had to help care for someone who has diabetes? [Choose one.]
    a. Yes [15.04%]
    b. No [84.96%]