Public and health professionals’ misconceptions about the dynamics of body weight gain/loss

Tarek Abdel-Hamid, a Felix Ankel, b Michele Battle-Fisher, c Bryan Gibson, d Gilberto Gonzalez-Parra, e Mohammad Jalali, f Kirsikka Kaipainen, g Nishan Kalupahana, h Ozge Karanfil, i Achla Marathe, j Brian Martinson, k Karma McKelvey, l Suptendra Nath Sarbadhikari, m Stephen Pintauro, n Patrick Poucheret, o* Nicolaas Pronk, p Ying Qian, q Edward Sazonov, r Kim Van Oorschot, s Akshay Venkitasubramanian t and Philip Murphy u

Tarek Abdel-Hamid
Department of Information Sciences, Naval Postgraduate School, Monterey, CA 93943, U.S.A.

Felix Ankel
University of Minnesota Medical School Emergency Department, Saint Paul, United States.

Michele Battle-Fisher
Wright State Boonshoft School of Medicine, Center for Global Health, Kettering, United States.

Bryan Gibson
University of Utah, VA Salt Lake City Healthcare System, Salt Lake City, United States.

Gilberto Gonzalez-Parra
Universidad de los Andes, Departamento de Cálculo, Merida, Venezuela.

Mohammad Jalali
Virginia Tech, Management Systems Engineering, Falls Church, United States.

Kirsikka Kaipainen VTT Technical Research Centre of Finland, VTT, ICT for Health, Tampere, Finland.

Nishan Kalupahana
University of Tennessee, UT Obesity Research Center and Department of Animal Science, Knoxville, Tennessee, United States.

Abstract

Human body energy storage operates as a stock-and-flow system with inflow (food intake) and outflow (energy expenditure). In spite of the ubiquity of stock-and-flow structures, evidence suggests that human beings fail to understand stock accumulation and rates of change, a difficulty called the stock–flow failure. This study examines the influence of health care training and cultural background in overcoming stock–flow failure. A standardized protocol assessed lay people’s and health care professionals’ ability to apply stock-and-flow reasoning to infer the dynamics of weight gain/loss during the holiday season (621 subjects from seven countries). Our results indicate that both types of subjects exhibited systematic errors indicative of use of erroneous heuristics. Indeed 76% of lay subjects and 71% of health care professionals failed to understand the simple dynamic impact of energy intake and energy expenditure on body weight. Stock–flow failure was found across cultures and was not improved by professional health training. The problem of stock–flow failure as a transcultural global issue with education and policy implications is discussed. Copyright © 2014 System Dynamics Society


Introduction

Regarding body weight, people are decision-makers managing a complex dynamic system: the human body. In addition to motivation, this task requires an understanding of how the human energy and weight regulation system

---

* Correspondence to: Patrick Poucheret, Laboratoire de Pharmaco logie et Physiopathologie Expérimentale, UMR 95 Quaisud, 15 Avenue Charles Flahault, 34093 Montpellier, FRANCE. E-mail: patrick.poucheret@univ-montp1.fr

Accepted by Laura J. Black; Received 17 April 2013; Revised 10 June 2014; Accepted 10 June 2014

System Dynamics Review
Published online in Wiley Online Library (wileyonlinelibrary.com) DOI: 10.1002/sdr.1517
works. In this paper, we report the results of a worldwide study assessing public and health care professionals’ (HCPs) understanding of human weight/energy regulation dynamics.

**Conceptual background**

Research in control theory and behavioral decision making suggests that effective control of a dynamic system requires two essential cognitive skills: (i) the ability to develop an adequate model of the system and (ii) the ability to run that model-inferring system behavior over time (Brehmer, 1990). For the operator to build an adequate model of a system, two levels of knowledge integration are required. First, structural knowledge establishes what the system variables are. Second, functional knowledge establishes how the variables are related to each other, as well as to the system’s external environment (Abdel-Hamid, 2009, 2012). Understanding and predictive skills, respectively referred to as system building and running, while interdependent, are nonetheless conceptually different (Einhorn and Hogarth, 1987). Research suggests that understanding system’s structure does not guarantee proper prediction of system behavior. Human reasoning appears to be inadequate to infer the behavioral implications of a system’s interactions over time (Sterman, 2002). These cognitive limitations may be a consequence of the “bounded rationality” that constrains human decision-making performance (Simon, 1970). It extends to even simple stock–flow (SF) systems, especially systems with nonlinearities, feedback and/or delays (Forrester, 1979; Senge, 1990; Diehl and Sterman, 1995; Gonzalez and Dutt, 2011). In this project we sought to answer the question: do these cognitive limitations apply to weight management and energy regulation?

Individual’s health management behaviors, such as dietary intake and physical activity, are associated with chronic disease later in life. Global public health issues related to dietary intake and physical inactivity include: obesity, cardiovascular diseases, diabetes, metabolic syndrome and some forms of cancer, among others. In this context, individuals’ predictive judgments are critical to understanding and implementing appropriate prophylactic and/or curative interventions (Hall et al., 2012). Predictive judgment allows individuals to prospectively anticipate resources to commit to a treatment and to retrospectively evaluate the use of allocated resources. In addition, regarding weight management, concordance between expected and obtained outcomes of intervention is associated with patient adherence to treatment and thereby with treatment efficacy (Dalle Grave et al., 2005).

Obesity, as an important and complex disease, stands “among the most misunderstood of human conditions” (Sontag, 1990; Shell, 2002). This applies to lay public, health professionals and academics (Casazza et al., 2013). Previous reports
have demonstrated that people misperceive their own weight status and its implications on health (Kuchler, 2003). Misconceptions may be related to misunderstanding the cause of a problem, its consequences, the timeline of its change, or in attributing specific symptoms to the underlying disease identity (Leventhal et al., 1980). Each of these dimensions has implications for people’s reaction when facing a disease threat, and all are relevant to obesity. This study focused primarily on timeline, which refers to disease time course, evolution and treatment, and on disease identity, which refers to symptomatology and associated complications.

The stock-and-flow operator model of human energy regulation

Most people intuitively understand that body weight varies as a function of the difference between energy intake and expenditure. The didactic analogy is the bathtub water level that rises and falls as a function of the rate of water flowing in or draining out (Figure 1) (Booth Sweeney and Sterman, 2000).

This simple SF structure is by no means unique to human energy regulation; rather, it characterizes many real-life systems and everyday tasks such as managing a checking account or a company’s inventory. It is because of this ubiquity that system dynamicists have long argued that understanding the relationship between stocks and flows is fundamental to understanding and managing a wide range of systems used in human activities (Gonzalez and Wong, 2012).

“Running” the operator’s model

In spite of the importance and ubiquity of stock and flow systems in the real world, research suggests that humans do not infer the dynamics of systems reliably (Booth Sweeney and Sterman, 2000). Indeed many problems, whether
individual, organizational or societal, are based on misunderstanding of the basic processes of SF dynamics (Cronin et al., 2009). The “SF failure” seems to be rooted in a “lack of human understanding of accumulation and rate of change” (Gonzalez and Wong, 2012, p. 3). This is often exacerbated by the application of inappropriate heuristics (Cronin et al., 2009). Regarding obesity, if SF failure does apply to human judgment about weight/energy regulation, it could undermine prevention and management efforts.

Study objectives

Traditional holiday season overeating, i.e. during North America’s Thanksgiving and Christmas, is associated with a 25–40 percent increase of caloric intake (Khare and Inman, 2009).

The aim of the present investigation was to conduct a worldwide study to assess public and health professionals’ understanding of stock-and-flow relationships that underlie the dynamics of weight gain/loss, through the case of holiday season overeating. For this purpose, we designed an experimental task to assess the degree to which people misjudge the dynamics of weight change associated with holiday overeating.

Method

Systems-Inspired Global Obesity Study (SIGOS)

Since obesity is recognized as a global health issue (Kelly et al., 2008), this study was conducted worldwide. Our primary research goal was to assess individuals’ understanding of bodyweight dynamics across different populations and cultures. Our secondary research goal was to investigate the extent to which professional health training affects performance. Our third research goal was to address the test proposed by Brunstein et al., who demonstrated that domain experience is not a strong indicator for overcoming SF failure. They concluded that follow-up studies should test results robustness through: (i) wider medical experience of participants; (ii) use of exactly the same method for all groups tested; and (iii) replication of the study in other domains of knowledge (Brunstein et al., 2010, p. 352). SIGOS study was built following Brunstein et al.’s recommendations.

Twenty researchers from health sciences and system dynamics fields, out of the 200 invited by Dr Abdel-Hamid via e-mail, accepted to collaborate. Each of these researchers drew samples from one or two distinct populations: lay people and health care professionals (HCP) from their local area. A total number of 621 subjects participated.
The experimental task

To assess lay persons’ and HCPs’ understanding of SF relationships pertaining to the dynamics of weight gain/loss, we created a typical SF task (Figure 1) that required subjects to determine how body weight (energy stock) varies in response to changes in caloric intake (the inflow) during the holiday season based on an overeating scenario. The exercise sheet (Figure 2a) shows a hypothetical subject at initial steady state of 150 lbs weight and 2000 kcal daily caloric intake and expenditure. The individual’s food intake increases linearly and progressively (by 25 percent), peaking at 2500 kcal/day on Thanksgiving and then declines back to 2000 kcal/day. This dietary pattern repeats during Christmas. To simplify the task, energy expenditure is assumed to remain constant throughout at 2000 kcal/day. The subjects’ task was to sketch on graph (b) their expectation of how body weight would change over the holiday season. The initial steady-state situation was pre-plotted. The subjects were given 10 minutes to complete the paper-based exercise. Subjects were not paid for participating (nor graded if students).
The solution and its characteristics

The task was designed to be simple enough to be correctly answered without the use of mathematics. To perform the task with success (Figure 3), subjects needed to apply simple rules governing the replenishment/depletion of a stock as a function of inflows/outflows. Rules that would lead to a correct solution are: (i) whenever the inflow (energy intake or EI) exceeds the outflow (energy expenditure or EE), body weight should rise; and (ii) whenever the inflow is equal to the outflow, body weight should hold steady. Body weight therefore depends on the relative magnitudes of the inflow and outflow; not whether the inflow is increasing or decreasing.

The experimental instrument was pilot tested on student subjects at the Naval Postgraduate School to identify typical responses. In addition, responses on comparable SF tasks published in the literature were catalogued. From these two inputs, a set of four “typical” responses was identified. These four typical sketches, together with two blank panels, were incorporated into a standardized data collection spreadsheet (Figure 4). Each researcher catalogued the number of responses in their sample that matched each of the four typical sketches. The two blank panels were provided so that researchers could record responses that did not fall into the four predefined types.4

Subjects

The SIGOS study guidelines specified a minimum sample size of 25 adult subjects and when possible from two distinct groups: lay subjects (mostly undergraduate university student) and health care professionals (Health Care Providers and medical students). Ultimately, 22 samples were drawn from 20 sites in seven countries: U.S.A., Venezuela, Norway, France, China, India and Sri Lanka. Two SIGOS researchers conducted two studies each, one with lay subjects and a second with health care providers; 621 subjects participated in the SIGOS study (71 percent U.S. subjects); 59 percent of the subjects were lay people and 41 percent HCP.
In order to explore our first question of whether experimental task performance varied with demographic characteristics, subjects’ age, sex, weight and height were requested. To protect subjects’ privacy, ID codes were assigned and used instead of names in coding and analysis. Institutional Review Board approval was obtained by each individual investigator as needed.

**Results**

A chi-square test revealed no significant difference in the distribution of HCP and lay persons between the U.S. and non-U.S. samples ($\chi^2 = 0.047$, d.f. = 1, $p = 0.83$). Demographic comparisons were conducted to assess possible differences between lay and HCP subjects (Tables 1 and 2). There was no significant difference in body mass index (BMI) ($t = 0.075$, d.f. = 619, $p = 0.94$) and there was a small but statistically significant difference in ages of 2.9 years ($t = 2.932$, d.f. = 604.7, $p = 0.003$). A chi-square test indicated a significantly larger proportion of females amongst HCP versus lay people ($\chi^2 = 15.86$, d.f. = 1, $p < 0.001$). To assess the subjects’ performance on the drawing task, we determined whether participants sketched the correct trajectory of body weight. We
graded the qualitative shape as correct if the participant sketched the right trajectory (Figure 3) and incorrect if they drew it differently in any way from the correct shape. Participants were not penalized for quantitatively incorrect drawing patterns.

Overall, subjects’ performance was low (Figure 5), with 26 percent of participants producing correct drawings. HCP (29 percent correct) did not perform significantly better ($\chi^2 = 2.54$, d.f. = 1, $p = 0.11$) than lay subjects (24 percent correct).

To gain further insight into individuals’ misunderstanding, we created three criteria to evaluate individuals’ drawings (a correct drawing would be correct for all three criteria).

**Table 1. Subjects’ demographic profiles**

<table>
<thead>
<tr>
<th>Professional status</th>
<th>Gender</th>
<th>Age</th>
<th>BMI</th>
</tr>
</thead>
<tbody>
<tr>
<td>HCP (n = 256)</td>
<td>Male (n = 314)</td>
<td>μ (σ)</td>
<td>μ (σ)</td>
</tr>
<tr>
<td>Lay (n = 365)</td>
<td>Female (n = 307)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(n)</td>
<td>%</td>
<td>(n)</td>
</tr>
<tr>
<td>U.S.A. (442)</td>
<td>41 (181)</td>
<td>59 (261)</td>
<td>49 (218)</td>
</tr>
<tr>
<td>France (50)</td>
<td>50 (25)</td>
<td>50 (25)</td>
<td>50 (25)</td>
</tr>
<tr>
<td>Norway (29)</td>
<td>0 (0)</td>
<td>100 (29)</td>
<td>59 (13)</td>
</tr>
<tr>
<td>China (25)</td>
<td>0 (0)</td>
<td>100 (25)</td>
<td>100 (25)</td>
</tr>
<tr>
<td>Venezuela (25)</td>
<td>0 (0)</td>
<td>100 (25)</td>
<td>28 (7)</td>
</tr>
<tr>
<td>India (25)</td>
<td>100 (25)</td>
<td>0 (0)</td>
<td>36 (9)</td>
</tr>
<tr>
<td>Sri Lanka (25)</td>
<td>100 (25)</td>
<td>0 (0)</td>
<td>52 (13)</td>
</tr>
</tbody>
</table>

**Table 2. HCP versus lay overall**

<table>
<thead>
<tr>
<th></th>
<th>HCP</th>
<th>Lay</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>151</td>
<td>156</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>105</td>
<td>209</td>
<td>$\chi^2 = 15.88$, d.f. = 1, $p &lt; 0.001$</td>
</tr>
<tr>
<td>Age</td>
<td>34.3</td>
<td>31.4</td>
<td>$t = 2.932$, d.f. = 604.7, $p = 0.003$</td>
</tr>
<tr>
<td>BMI</td>
<td>24.3</td>
<td>24.2</td>
<td>$t = 0.075$, d.f. = 619, $p = 0.94$</td>
</tr>
</tbody>
</table>

Fig. 5. Fraction of subjects producing correct sketch
Drawings were scored “1” if correct for each criterion and “0” if incorrect. The majority of subjects drew trajectories that violated one or more of the three criteria (Table 3):

- 44% correctly drew weight rising when food intake was falling (after peaks III-1 and III-2);
- 50% correctly kept weight elevated as energy intake drops between the two peaks (III-2);
- only 36% placed the peak of final body weight (III-3) at the right time.

Again, on the three criteria, HCP did not perform significantly better than lay people (Figure 6).

To assess performance on all three criteria, we combined the three individual scores into a normalized composite score; a correct criterion was scored “1”,
while violation of any criterion scored “0”. As expected, this alternative form of scoring subjects’ responses produced results (Table 4) similar to the graph-based analysis (Figure 5). HCP incorporated all three features of a correct graph somewhat more often than lay people. There was a little more variation in HCP responses ($\sigma_{\text{lay}} = 0.42$ vs. $\sigma_{\text{HCP}} = 0.45$), but differences in performance were limited (28 percent for HCP vs. 23 percent for lay) and not significant ($\chi^2 = 2.01$, d.f. = 1, $p = 0.16$). Using the normalized composite indicator, variations on a logistic regression were then run to test whether age ($\mu = 32.6, \sigma = 12.6$), BMI ($\mu = 24.2, \sigma = 4.6$), gender (F = 307, M = 314), or HCP status predicted a subject’s success in producing the correct answer. Neither age ($p = 0.94$) nor BMI ($p = 0.51$) was a significant predictor of success (Table 5). HCP status, when considered on its own (Model 4), was not a significant predictor of ability to produce the correct graph. However, we observed a marginally significant ($p = 0.052$) gender difference in likelihood of success between lay and HCP, with females performing worse than males in this sample. To assess for confounding, we tested for interaction between gender and professional status (Model 2) and then subsetted the

| Normalized total score (NTS) | | |
|--------------------------------|-----|
| NTS score for all subjects | 0.25 |
| NTS for all HCP subjects | 0.28 |
| NTS for non-HCP subjects | 0.23 |

| Table 5. Logistic regression models | | | | |
|-------------------------------------|-----|
| Correct representation of weight trajectory | | | |
| (1) | (2) | (3) | (4) | |
| Age | $-0.003$ | $0.018$ | $-0.592$ | $0.018$ |
| | (0.008) | (0.021) | (0.197) | (0.021) |
| BMI | $0.018$ | $-0.490$ | $-0.628$ | $0.504$ |
| | (0.021) | (0.261) | (0.193) | (0.256) |
| Gender (female = 1) | $-0.592$ | $-0.490$ | $-0.628$ | $0.018$ |
| | *** | * | *** | |
| | (0.197) | (0.261) | (0.193) | (0.256) |
| Professional (HCP = 1) | $0.372$ | $0.504$ | $0.373$ | $0.265$ |
| | * | ** | * | |
| | (0.193) | (0.256) | (0.192) | (0.187) |
| Gender: professional | $-0.297$ | $-0.297$ | $-0.981$ | $-1.223$ |
| | (0.386) | (0.386) | (0.142) | (0.125) |
| Constant | $-1.348$ | $-1.030$ | $-0.981$ | $-1.223$ |
| | *** | *** | *** | *** |
| | (0.539) | (0.157) | (0.142) | (0.125) |
| ObservationsAkaike inf. crit. | 621 | 621 | 621 | 621 |
| | 691.767 | 690.259 | 688.851 | 697.661 |
data by gender and reran the logistic regressions on the two subsets (not shown). We observed no significant interaction between gender and professional status ($p = 0.441$). Additionally, in the logistic regression models with males only, we observed that, among males, there was a significant difference in the likelihood of success between lay and HCPs ($p = 0.0474$). Interestingly, a similar gender effect was observed as in the original Booth Sweeney and Sterman study (2000); in their study, males were also marginally more likely to be successful than females.

Figure 7 and Table 6a show frequencies of subjects’ graphed responses organized into three categories: correct answer, use of correlation heuristic and otherwise incorrect answers. Correlation heuristic refers to responses that match the pattern of graph 1 in Figure 4, in which body weight matches the pattern of food intake. Use of correlation heuristic was the most frequent type of incorrect answer: 36 percent of participants produced this type of graph. There was no association between the graph type drawn and the subject status as HCP (Table 6b) ($\chi^2 = 2.34$, d.f. = 2, $p = 0.31$).

In order to assess between-country differences, we compared average task performances plotted against obesity rate for each country in Figure 8. The scatter plot provides a “performance × importance” analysis (IPA) comparison. IPA is a diagnostic and strategic planning approach to assess the consistent allocation of resources, a key determinant of performance, to activities deemed to be important for a society (Eskildsen and Kristensen, 2006).

![Fig. 7. Observed responses by graph pattern (graph 1 represents correlation heuristic, graph 4 represents correct pattern, graph 5 represents all other patterns)](image)

Table 6. a) Patterns identified overall. b) Patterns identified by professional status

<table>
<thead>
<tr>
<th></th>
<th>Correct</th>
<th>Incorrect</th>
<th>Correlation heuristic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Counts</td>
<td>161</td>
<td>235</td>
<td>225</td>
</tr>
<tr>
<td>(%)</td>
<td>25.93%</td>
<td>37.84%</td>
<td>36.23%</td>
</tr>
<tr>
<td>HCP counts</td>
<td>74</td>
<td>96</td>
<td>86</td>
</tr>
<tr>
<td>(%)</td>
<td>28.9%</td>
<td>37.5%</td>
<td>33.6%</td>
</tr>
<tr>
<td>Lay counts</td>
<td>87</td>
<td>139</td>
<td>139</td>
</tr>
<tr>
<td>(%)</td>
<td>23.8%</td>
<td>38.1%</td>
<td>38.1%</td>
</tr>
</tbody>
</table>
While preliminary, IPA results highlight two interesting questions: (i) is there a causal link between poor performance in assessing weight gain/loss dynamics and effective personal weight management? and (ii) why is the U.S.A. positioned at the lower-right quadrant on the IPA grid?

**Discussion**

The SIGOS study builds upon a long stream of research investigating decision-making performance on SF tasks (Brehmer, 1992; Sterman, 2000; Sterman and Booth, 2002; Maani and Maharaj, 2004; Li and Maani, 2011). This study demonstrated that individuals’ problems in understanding SF systems extend to their understanding of body weight regulation. This finding has significant implications for public health messaging, health education and health policy. In addition, this study pilot tested a novel social-networking-inspired model for research collaboration and demonstrated its feasibility.

**Key findings and implications**

In this study we presented highly educated subjects (postgraduate) with a simple task to test their ability to reason about basic SF dynamics relating to weight gain/loss. Overall, both lay and HCP subjects performed poorly. This study demonstrates for the first time that the health issue of regulation of diets to control weight is no exception to SF failure. More specifically, we found that individuals failed to grasp that quantity of a stock rises when the inflow exceeds the outflow, and intuitively assume that the output of a system is positively correlated with its inputs. This cognitive error, known as “the correlation heuristic”, is well established (Booth Sweeney and Sterman, 2000; Atkins et al., 2002; Pala and Vennix, 2005; Cronin et al., 2009). Previous investigations of this SF failure have ruled out potential confounders, including lack of...
contextual knowledge, inadequate motivation, inability to interpret graphics, confusing information displays, cognitive effort and computational complexity (Cronin and Gonzalez, 2007; Cronin et al., 2009). Therefore, SF failure appears to be a pervasive deficiency in human reasoning and not an artifact of the task (Cronin et al., 2009). Our results corroborate prior findings from SF experimental studies covering a wide variety of task settings (Booth Sweeney and Sterman, 2000).

Two aspects of the experimental task that many participants, both lay and HCP, found particularly counterintuitive deserve some discussion. First, there was a widespread tendency among subjects to match the trajectories of body weight to the inflow pattern. This use of a correlation heuristic illustrates a basic misunderstanding of the difference between what causes a flow to move in a certain direction (increasing/decreasing) and what causes the associated stock to rise or fall. The correlation heuristic is sometimes appropriate to apply (e.g. illness is highly correlated with the consumption of toxic mushrooms), but in systems with SF structures it can lead to systematic errors, as it does in this case (Cronin et al., 2009). Second, many subjects failed to properly account for the inertia and memory inherent to the process of stock accumulation. After the Christmas feast, body weight does not drop to its original level, even as food intake drops back to pre-holiday level. Rather, body weight remains at its maximum level. Indeed, at any point in time, the amount of energy in the body (the stock) reflects the cumulative effect, over time, of the net inflows over outflows, not merely the instantaneous flow status at a particular time point. Therefore, stocks are said to provide systems with inertia because they establish a “memory” of all past events in the system. Failure to recognize the stock-related inertia/memory in personal health may be a “self-serving” and presumably pervasive misconception that has potentially serious implications in judging and managing health risks (Hall et al., 2012). For example, weight gain and lack of compensation after holiday season overeating have been previously reported (Brownell and Horgen, 2004; Yanovski et al., 2000; Parker-Pope, 2005).

Our study also revealed a particularly disturbing observation: HCPs did not perform better than lay subjects. Many HCPs drew trajectories that violated even basic physical principles such as the conservation of energy/matter, i.e. failing to realize that the quantity in a stock must increase if the inflow exceeds outflow, even if inflow is decreasing. This observation raises the question of the impact of domain experience in overcoming SF failure. Brunstein et al. (2010) explored lay persons’ and medical students’ SF reasoning performance on medical and non-medical accumulation problems, using various task displays. These authors showed that domain experience (duration of medical education and age) “is not a strong indicator for SF performance”. In SIGOS, practicing HCPs from different countries and medical educational systems performed similarly to lay subjects and medical students in terms of SF failure, suggesting that, in addition to duration of education, daily medical practice does not improve SF performance. Therefore SIGOS results corroborate the conclusions of prior work on SF failure and extend them to the domain of weight regulation.
Assuming our results are replicated, they raise serious concerns. Fundamentally human beings cannot effectively manage what they do not understand. Therefore, misconception of the SF dynamics underlying body weight will lead to erroneous inferences about their health-associated risks. This could severely undermine obesity prevention and management efforts. In consequence, both lay (general public and undergraduate) and HCP (medical graduate and practitioners) might benefit from developing an intuitive appreciation of the basics of SF dynamics in order to use an appropriate heuristic. Linking these concepts to human bioenergetics should promote a paradigm shift and create a common consensual cognitive framework. For lay people it would clear up confusion, favor new habits in weight management and sensitize them to obesity prevention programs. For HCPs, it would improve understanding of pathology and therapeutic strategies as well as mutual understanding with patients, thereby potentially improving medical care compliance and efficacy. Poor performance by HCP on SF tasks underscores the need to incorporate systems thinking into health education curricula. One potential starting point could be the use of teaching-pictures recently shown to simplify resolution of accumulation problems (Brockhaus et al., 2013). This might help individuals to develop intuition for dynamics. Overall, the social-networking-inspired model for research collaboration has pros (cost/time effectiveness, transdisciplinarity and worldwide dimension) and cons (decontextualized interactions, terminology issues resulting from transdisciplinarity). This research collaboration model requires a high level of commitment but generates reciprocal scientific emulation.

Future research directions

Because obesity rates differ widely among populations and/or cultures, the levels of public discourse and public health “educational” interventions differ widely as well. This raises a number of policy questions: what effect, if any, do public awareness and public health education have on enhancing people’s understanding of body weight/energy regulation and of its dynamics? And does it matter? Unraveling the effect of the causal link, if any, between understandings of SF dynamics related to weight gain/loss and effective personal weight management is an experimentation challenge. Since obesity is a complex multifactorial condition, it is unlikely that understanding the dynamics of personal weight/energy regulation is a sufficient condition on its own for effective weight management. But is it a necessary condition? IPA results suggest it may well be a promising research endeavor with potentially significant policy and health implications. This point is perhaps best illustrated by Sri Lanka’s intriguing results, whose HCPs achieved the highest performance, with 64 percent correct responses at a low 5 percent obesity rate. The IPA raises a second critical policy question: why is the U.S.A. in the lower-right quadrant of the grid, despite its government’s committed educational efforts about obesity and guidelines to promote...
healthy weight? For the U.S.A. not to be in the upper-left quadrant, high performance and low obesity rate may be related to significant delays encountered in translating public understanding into behavioral changes and in creating supportive environments (physical and economical). For the U.S.A. not to be in the upper-right quadrant (high performance and high obesity rate) is surprising and might be of concern. To understand whether this is a public policy failure, and to propose alternatives, is a legitimate concern that needs to be addressed.

In conclusion, altogether, our findings suggest that SF failure is a global problem rather than a problem unique to a particular culture. Therefore it might be possible to envisage a common solution. One potential intervention would be to provide system-dynamics-based insights about health issues such as obesity and associated complications to individuals and their health care providers. We believe this could contribute to improved human health worldwide.

Notes

1. Non-U.S. participants were asked to substitute the Thanksgiving/Christmas designations with their own pair of local holidays (a pair of eating-type holidays that are approximately 1 month apart) to ensure the cover story fits the cultural context of their study.
2. In reality, even if physical activity is unchanged, a person’s total daily energy expenditure (EE) changes if body weight and composition change in response to overeating. For example, weight gain often triggers metabolic adaptations that increase resting energy expenditure (REE). However, changes in total EE induced by 4–8 days of overeating (as per our scenario) would be quite small.
3. A note to the reader: before proceeding to the next section, we suggest you take a few minutes to sketch your own solution.
4. After the study was conducted, experimental subjects were referred to an online implementation of the task for the opportunity to “play” it again online and/or to access an interactive tutorial explaining the correct solution. The online version remains available for public access on: http://forio.com/simulate/tkabdellhs-irsl-u-forio/overview/.
6. The simple graphs had five points and represented a 5-minute time period, whereas the complex graphs had 21 points and represented a 21-minute time period. The shapes of the graphs were identical.

Acknowledgements

We thank Dr Alba Salas Paredes, Facultad de Farmacia y Bioanálisis, Universidad de Los Andes, Mérida, Venezuela, for his contribution in the acquisition of data.
and Dr. Karen T. Cuenco, University of Pittsburgh, for her input and comments on the statistical analysis. Financial support (Prof. Marathe) was provided through NIH MIDAS Grant 2U01GM070694-09 and 3U01FM070694-09S1; DTRA Grant HDTRA1-11-1-0016. We thank anonymous reviewers for their valuable and constructive comments and are grateful to SDR editors and Laura Black for their insightful advice.

References


