Do feedback diagrams promote learning in macroeconomics?

I. David Wheat

University of Bergen,
Fosswinkelsgate 6, N-5007 Bergen, Norway
E-mail: david.wheat@uib.no

Abstract: This article describes the value added by a stock-and-flow feedback diagram to text-only instruction in macroeconomics. The experiment was motivated by a prior study in which the use of graphs to teach macroeconomics was no more effective than verbal instruction alone. Here, in contrast, students using the feedback diagram showed more improvement in post-test scores than those who received only narrative instruction.

Keywords: conceptual change; dynamics; education; feedback; experiment; learning; macroeconomics; mental model; meta-analysis; stock-and-flow; value added.


Biographical notes: I. David Wheat is an Associate Professor of System Dynamics at the University of Bergen in Norway, a Visiting Professor of Economics at ISM University in Lithuania, and an Adjunct Professor of Economics at Virginia Western Community College in the USA. He recently served as the President of the Economics Chapter of the International System Dynamics Society.

1 Introduction

Would students get a better grasp of macroeconomic concepts if lectures, textbooks, and other instructional materials included diagrams that illustrate feedback processes in the economy? The answer may seem self-evident, given the conventional wisdom regarding illustration of abstract concepts. The popular claim that a picture is worth a thousand words is grounded in the education psychology literature (cf., Standing, 1973). And evidence that the human brain has a remarkable capacity for long-term storage and retrieval of visual images (Bahrick et al., 1976) may be reason enough that textbook presentations of abstract concepts are routinely supplemented by illustrations. In addition, research suggests that pictures and diagrams facilitate learning at levels deeper than mere retention and recall, at least for some students. For example, findings by Shaw (2000) and Wolfe (2001) indicate that comprehension of information is fostered by visualisation of that information, at least for those aptly described as ‘visual learners’.
In undergraduate economics lectures and texts, this ‘illustrate the concept’ strategy relies on the ubiquitous and multifarious statics graph (Kennedy, 2000; Cohn et al., 2001), with the tacit assumption that such graphs facilitate learning. For example, Mankiw’s (2004, chapter 20) discussion of how “we analyze fluctuations in the economy as a whole” is supplemented by nine graphs depicting various static representations of aggregate demand and aggregate supply. According to Cohn et al. (2001), the number of graphs in popular modern economics textbooks ranges between 200 and 400, the midpoint of which is at least ten times more than early 20th century texts.¹

To evaluate the instructional efficacy of graphs, Cohn et al. (2001) conducted two experiments. In both, they expected “that use of graphs in a lecture would benefit students, especially in economics classes where the material presented is extremely abstract.” In one experiment, however, they found a graph-supplemented lecture to be no more effective than verbal instruction alone; in another, verbal instruction was more effective (Cohn et al, 2001). By questioning the pedagogical value added by graphs used in macroeconomics textbooks and lectures, they challenged a broader conventional wisdom. This paper extends their line of questioning by testing whether an alternative illustration method would have pedagogical value in a macroeconomics course.

Since publication of Cohn’s experimental results, there have been few attempts to replicate or otherwise explore the implications of their findings. In one replication attempt that ostensibly contradicts Cohn, introductory economics students who used graphs scored higher on a quiz than those who did not (Sedaie, 2004). In that experiment, however, the students using graphs were self-selected and, despite including some control variables, it was unclear whether the difference was due to the graphs or to characteristics of students who chose to use them. Another response to the Cohn experiment comes from Hansen et al. (2002, p.467), who recommend “that instructors develop graph-free strategies for teaching most concepts.”

This paper discusses my experiment that was inspired by Cohn’s original design but offers an alternative teaching strategy. As in the Cohn experiment, two groups of students received identical narrative instruction on a macroeconomics topic, with one having a supplemental visual aid. Unlike the Cohn experiment with its reliance on graphs for a visual aid, my experiment tested the value added by a stock-and-flow feedback diagram, one of the central constructs in the feedback method of teaching macroeconomics.

2 Improving mental models

My hypothesis is that a feedback representation of macroeconomic processes is pedagogically useful for improving undergraduates’ mental models of the macroeconomy. A desirable mental model is one that enables a dynamic, rather than static, conception of the economy and focuses on the process of change rather than equilibrium. Higher order mathematics is usually presumed to be a prerequisite for understanding dynamics in a complex system, as evidenced by the mandatory math training for upper level economics majors. Yet, the potential for students to understand economic dynamics prior to receiving advanced mathematics instruction is suggested by my experience teaching economic feedback processes to undergraduates and also by the results of three additional experiments (Wheat, 2007).² The experiment described in this paper provides additional support for my hypothesis.
Do feedback diagrams promote learning in macroeconomics?

2.1  Mental models

The mental model concept is rooted in cognitive psychology and stems from Craik’s (1943) suggestion that the human mind can construct ‘small-scale models’ of reality for the purpose of understanding, explaining, or anticipating events in the real world. Nearly, four decades passed, however, before the concept emerged as an important topic in the psychology literature and a useful cognitive construct. According to Rouse and Morris (1986, p.351), “mental models are the mechanisms whereby humans are able to generate descriptions of system purpose and form, explanations of system functioning and observed system states, and predictions of future system states.” And Seel (2001, p.407) points out that both cognitive psychologists and educational psychologists consider mental models as “qualitative mental representations which are developed by subjects on the basis of their available world knowledge aiming at solving problems or acquiring competence in a specific domain.” Although Bruner (1960) did not use the term ‘mental model’, he emphasised the pedagogical significance of placing instructional details into a ‘structured pattern’ lest the details be forgotten or their meaning slip away.

And the role of mental models in shaping perceptions of complex systems has long been emphasised in the system dynamics literature (cf. Senge, 1990; Doyle et al., 2001). For example, in an essay on the counterintuitive behaviour of social systems, Forrester (1971, p.3) stressed that “the mental image of the world around us that we carry in our heads is a model. One does not have a city or a government, or a country in his head. He has only selected concepts and relationships, which he uses to represent the real system.” In his discussion of mental model construction and renovation in an instructional setting, Seel (2001, p.411) emphasised that learners must be provided “an appropriate conceptual model of the facts to be causally explained.” System dynamics modellers have developed conceptual tools to facilitate communication between model builder and model user regarding systemic structure and behaviour. One such tool is the simple feedback diagram, either in stock-and-flow format (such as the GDP model used in this experiment) or simple word-and-arrow diagrams that illustrate feedback loops (cf., Wheat, 2009b). The potential for feedback diagrams to add instructional value was suggested by Forrester’s (1994) reference to system dynamics as a “framework into which facts can be placed [so that] learning becomes more relevant and meaningful.” Forrester’s framework is Bruner’s (1960) structured pattern.

At any point in time, students’ mental models are imperfect approximations of their perceptions of reality. Moreover, a change in that perception is a prerequisite for a change in the mental model. Indeed, one could define teaching about real-world systems as an attempt to change students’ mental models by improving their perception of how those systems work. Since a mental model is ‘relatively enduring’ [Doyle et al., (2001), p.5], the requisite perceptual change occurs gradually, by way of a self-adjusting learning cycle or counteracting feedback learning loop (Kolb, 1984; Sterman, 2000) that gradually updates prior perceptions and the associated mental model.

However, Doyle et al. (2001) emphasise the general difficulty of improving perceptions of the structure of dynamic systems and thereby enabling reasonable expectations of the behaviour of such systems. These difficulties stem from a lack of training in stock-and-flow thinking; the tendency to think in terms of one-way causality instead of feedback processes; long and variable delays that make difficult the association of today’s cause with tomorrow’s effect; non-linear functional relationships that abound in social systems and make extrapolation of trends an unreliable forecasting method; and
the inherent cognitive limitations on human capacity to process multiple relationships and mentally simulate the behaviour of complex systems.4

The stock-and-flow feedback process that is the conceptual foundation for this experiment does not address all of the difficulties inherent in renovating mental models of dynamic economic systems. Hopefully, it is a step in the right direction. The underlying premise is that when teaching about a complex dynamic system, visual aids that clarify the structure responsible for processes of change over time may facilitate desired adjustments in students’ mental models. This is the pedagogical purpose of the stock-and-flow feedback diagram used in the experiment.

2.2 Stocks, flows, and feedback loops

System dynamics distinguishes between a system’s condition and changes in that condition, conceptualised as stocks and flows, respectively. The importance of such a distinction in economics is evident to instructors who witness student confusion when discussing such topics as debt vs. deficit, wealth vs. income, and savings vs. saving.5

The conceptual building blocks for the feedback method of teaching macroeconomics are stocks, flows, and feedback loops. These concepts are central to the principle that dynamic behaviour occurs as flows accumulate in stocks (e.g., when a production inflow adds to a stock of inventories, and a sales outflow subtracts from inventories). A feedback system is one in which the flows change endogenously due to the subsequent influence of the stocks (e.g., when future production is influenced by an assessment of the adequacy of current inventories in light of expected sales).6

Feedback processes are either reinforcing or counteracting in their effects. The familiar compounding process involving interest and principal in a bank account illustrates the power of a self-reinforcing (or ‘positive’) feedback loop. The classic example of a counteracting (or ‘negative’) feedback mechanism is the thermostat that regulates room temperature, and this example is also a reminder that negative feedback loops perform a self-adjusting, goal-seeking role (explicitly or implicitly).

From the perspective of the system dynamics paradigm, stock-and-flow feedback thinking about the macroeconomy can facilitate learning about aggregate production (e.g., GDP) as part of an endogenous feedback process. The value of this way of thinking becomes clear when students perceive its applicability to such diverse macrodynamic topics as countercyclical monetary and fiscal policies (negative feedback loops designed to counteract upward spirals in inflation or unemployment); housing price bubbles that burst (when a positive speculative demand loop is eventually counteracted by a negative consumer demand loop); and economic growth and business cycles (respectively influenced by the interaction of positive and negative feedback loops over time).

Of course, the feedback perspective is not unique to system dynamics. It permeates the history of economic thought, beginning perhaps with Cantillon’s (1755) hypothesis that rising wages encouraged 18th century British workers to have larger families, which contributed (years later) to a labour supply that outstripped demand and, ultimately, put downward pressure on wages. The 19th century list of feedback thinkers in economics includes John Stuart Mill (1848), who described the self-reinforcing process involving speculative demand and inflationary commodity markets. Samuelson’s (1939) multiplier accelerator model, which legitimised the endogenous perspective for many 20th century economists, can be translated from the original algebra to a diagram consisting of two reinforcing loops and a single counteracting loop (Richardson, 1991). More recently,
Do feedback diagrams promote learning in macroeconomics?

Gjerstad and Smith (2009) explained the USA housing price bubble in terms that could be translated into three (weak) counteracting feedback loops and four (strong) reinforcing loops (Wheat, 2009a).


The generic stock-and-flow diagram in Figure 1 displays and summarises the relationships among the key concepts introduced above.

Figure 1  Generic stock-and-flow diagram

To reiterate, the stock is an accumulation of material (e.g., inventories) or information (e.g., sales data that are collected and analysed prior to decisions about future production). The net flow is the rate of change in the stock. The feedback loop transmits information about the state of the system from the stock to the decision rules that govern the flow. The flow, in turn, updates the stock. In the experiment described below, these building blocks are used to construct a simple system of stocks and their associated flows, in the form of a visual aid for students learning about GDP.

3  The experiment

Two versions of the experiment were conducted, one in Massachusetts (site 1) and the other in Virginia (site 2). Meta-analysis statistical methods were used to aggregate and analyse data generated by 85 student participants (46 juniors and seniors from Harvard Public Schools in Massachusetts and 39 mostly first-year students from Virginia Western and Dabney Lancaster community colleges in Virginia). None of the participants had prior experience with system dynamics stock-and-flow concepts or diagrams. However, 85% were enrolled in a one-semester economics course and had received some macro instruction; thus, at least that many had some economics background.

Students were randomly assigned to a control group or experimental group. Only those in the experimental group saw the stock-and-flow visual aids. The explicit learning goals for both groups included improving students’ ability to define GDP and clarify its meaning, describe alternative ways of measuring GDP, and explain how GDP fits in the ‘bigger picture’ of a national economy. Implicit in the instructional content for the experimental group was an additional learning goal; namely, to improve students’ awareness of the dynamic process by which GDP changes.
Operational indicators of learning included answers to a set of multiple choice questions. Prior to instruction, all students took a pre-test to establish a benchmark for measuring improvement after instruction.

3.1 Instructional procedures and content

During instruction, students in both groups worked alone on computers, using the STELLA software story-telling feature to advance the slides of the GDP story. The story read by the control group contained only textual information about the meaning of GDP, its measurement, and its placement in an overall macroeconomy. We refer to this as the text-only instructional method.

Students in the experimental group read the same textual information, but it was accompanied by a stock-and-flow feedback diagram that revealed the structure of a simple economy in a manner designed to complement the narrative. This is the feedback instructional method. Figure 2 displays the completed version of the stock-and-flow diagram. During the experiment, the diagram was revealed to the students in a series of small steps, each accompanied by narrative annotation.

Figure 2 This completed stock-and-flow feedback diagram was revealed in stages for students in the experimental group during the instructional phase of the experiment (see online version for colours)

3.2 Experimental hypothesis

The effectiveness of the instructional methods is defined in terms of post-test score improvement relative to a benchmark established by a pre-test. Improvement is measured as a student’s normalised percentage gain (NPG) in test score. The NPG is the difference
Do feedback diagrams promote learning in macroeconomics?

between individual pre- and post-test scores, expressed as a percentage of the maximum possible improvement in the pre-test score.

\[ \text{NPG} = \frac{100 \times (\text{posttest score} - \text{pretest score})}{(100 - \text{pretest score})} \]

The denominator in the NPG equation is the learning ‘gap’ that needs to be closed, based on the pre-test. The quantity in parentheses in the numerator is the absolute gain. Thus, the NPG is the percentage of the gap that is closed after the instruction.\(^{12}\) Let \( G_{TS} \) and \( G_{FS} \) refer to the mean NPG for the text-only and feedback instructional methods, respectively. The null hypothesis, then, is

\[ H_0 : G_{TS} = G_{FS} \]

which implies no difference in the learning effects of the two instructional methods.

3.3 Meta-analysis

The use of slightly different test instruments in Massachusetts and Virginia precludes traditional pooling of the data for analysis. In this case, meta-analysis is more appropriate for aggregating the results.\(^{13}\)

“Meta-analysis is a body of statistical methods that have been found useful in reviewing and evaluating research results. If a number of independent studies have been conducted on a particular subject, using different data sets and methods, then combining their results can furnish more insight and greater explanatory power than the mere listing of individual results. ... High returns in the advancement of empirical understanding await economic researchers willing to develop and apply meta-analysis.” [Stanley, (2001), pp.131–133]

As typically practiced, meta-analysis involves statistical analysis of a large number of individual studies of a particular research question over a long-time period, with the purpose of systematically and objectively aggregating and evaluating the weight of the cumulative evidence.\(^ {14}\)

The meta-analysis utilised here involves just two studies spanning less than six months. Unlike most meta-analyses, our purpose was not to strengthen our understanding of studies done by others; rather our more modest goal was to strengthen the interpretation of our own studies which are more similar in design than most pairs of studies in the meta-analysis research literature. This meta-analysis treats the two sites’ instructional outcomes as a sample of size two. For that sample, a weighted average of the results was calculated, based on the standardised mean and variance and the sample size at each site. That produced an aggregate point estimate of the standardised difference in means for the text-only method and the feedback method. Analysis of variance techniques established probability levels and confidence intervals (cf., Borenstein et al., 2005).

4 Results

At both sites, the student group having the feedback diagram as a visual aid outperformed the group with text-only instruction.
4.1 Data

The data generated by the experiment are displayed graphically in Figure 3. At site 1, the mean NPG within the feedback group (GF1) was 32%, compared to 16% for the text-only group (GT1). The corresponding difference at site 2 was 49% for GF2 and 18% for GT2.

**Figure 3** Average improvement was greater in feedback groups

4.2 Analysis

Meta-analysis incorporates the results in Figure 3, along with the variance and sample size for each group at both sites, and derives an aggregate weighted average standardised difference in means between the text-only group and the feedback group. It is that difference that is evaluated for statistical significance. In this case, the difference is 41.0% with a probability of less than 6% that the null hypothesis is true.

Nevertheless, the statistical significance level is borderline – slightly above the 5% cut-off point usually deemed necessary to reject the null hypothesis. Below, we explore issues that are relevant to interpreting the results, with the goal of identifying the potential and direction of any bias. Specifically, we look for potential bias arising from sample selection, instructional content, and time-on-task.

4.3 Bias? Which way? Implications?

4.3.1 Sample selection

At least 85 participants had prior economics education, while none had system dynamics training. That could affect the outcome in two ways, both of which would dilute the effect of the feedback diagram and bias the results toward the null hypothesis. First, familiarity with economic terms and concepts, such as GDP, could have generated pre-test scores for both groups that would be higher than expected from students without any economics education. Higher baseline pre-test scores would reduce somewhat the
opportunities for improvement as a result of any instruction, and would diminish the potential for the more effective mode of instruction to display differential effects. Secondly, the use of familiar economic terminology in the text-only instruction could have had a reinforcing effect on students with prior economics education. That would have inflated the impact of the text-only method, making the beneficial impact of the feedback diagram harder to detect. In short, the background of the participants may have biased the results in favour of the null hypotheses – contrary to the actual results. In the absence of such bias, the confidence level achieved by the results might have been even higher.

4.3.2 Instructional content

A comparison of the instructional content for the two methods (available from the author) should confirm that the students in both groups received the same textual information about GDP. The only distinction was that the experimental group also received a stock-and-flow feedback diagram (with minimal additional annotation). Moreover, a careful reading of the instructional content reveals that the text-only instruction contained numerous references to stocks, flows, delays, and even the bathtub analogy, a staple of simple system dynamics exposition. In other words, the students in the control group received textual instruction that was replete with system dynamics concepts. The similarity of the two methods, therefore, was not limited to information content. Both also had a system dynamics conceptual foundation. The only difference was that the control group lacked the visual reinforcement of the stock-and-flow feedback diagram. The instructional content, therefore, could not be considered biased in favour of the diagram.

4.3.3 Time-on-task

A criticism could be directed at the relative time requirements of the two instructional methods. Most students in the control group completed the textual instruction and post-test in about 15 minutes. The students in the experimental group needed about twice as long to complete their task. Arguably, the additional time devoted to studying the feedback method could be partially responsible for the post-test score improvement. However, whenever diagrams are supplementary to textual information that constitutes extra instructional material for students. Studying the extra visual material requires extra time. If the textual material is the same – as the experiment required – the overall time requirements will necessarily increase. In real instructional settings, the choice is rarely between having or not having visual aids. The issue is what kind of visual aid, and the results of this experiment suggest that alternatives to graphs deserve consideration.

5 Conclusions

This paper was motivated by Cohn’s (2001) comparison of graphs and narrative instruction, and it is an indirect comparison of feedback diagramming methods and conventional graphical instruction. The results here suggest that feedback diagramming adds value to mere narrative instruction by facilitating perception of systemic structure and its attendant behaviour.
Despite the reasonable strong statistical support for the findings reported here, there is a need to upgrade the design of this experiment in ways that justify even higher levels of confidence. In particular, participants should have no prior exposure to economics, and better field-testing of questions is necessary to reduce the variance in mean response measurements, a prerequisite for raising statistical confidence.

Beyond these design issues, the scope of the research should be widened to learn more about how students’ form perceptions of the structure and behaviour of economic systems, how that perception formation process is influenced by instructional methods, and ultimately – how students’ mental models of economic systems can be most effectively and efficiently improved.

Acknowledgements

The author would like to thank two anonymous referees and the editor for insightful and constructive suggestions.

References

Do feedback diagrams promote learning in macroeconomics?


Notes
1 According to Cohn et al. (2001), Mankiw’s (1997) text had 197 graphs and Parkin’s (2000) text contained 417 graphs. In contrast, Taussig’s (1924) had only 12 graphs.
2 The author engages students in feedback diagramming and computer simulation in a macroeconomics principles course delivered via the Internet to students enrolled at Virginia Western Community College in the USA.
3 The first two books on mental models were Gentner and Stevens (1983) and Johnson-Laird (1983).
4 On the issue of misperception of feedback, see Sterman (1989) and Moxnes (1998).
Do feedback diagrams promote learning in macroeconomics?

In general, stock-and-flow distinctions are often lost on a public steeped in an educational system that fails to teach such concepts, as the scientific literature attests (cf., Cronin et al., 2009). A recent piece in the popular press – the December 2009 ‘big idea’ in *National Geographic* – represents a significant opportunity for stock-and-flow thinking to improve readers’ mental models about atmospheric CO₂ concentrations and climate change. http://ngm.nationalgeographic.com/big-idea/05/carbon-bath

Standard system dynamics references include Forrester (1961) and Sterman (2000).

For readers familiar with calculus, the ‘rate of change’ concept refers to differential equations. The net flow is the first derivative of the stock and the stock is the integral of the flow. In a time-series graph, displaying year-to-year trends in the value of a stock (e.g., product inventory or employment level or money supply), the slope of the stock value between two years is equal to the net flow into (or out of) the stock during that time period.

I am grateful to Ralph Sass and Larry Weathers for administering the experiment at Dabney Lancaster Community College and Harvard Public Schools, respectively.

The test questions and answers are available from the author upon request.

**STELLA** is user-friendly software for modelling dynamic systems with the building blocks of stocks, flows, and feedback loops. The name is a registered trademark of isee systems, inc. (http://www.iseesystems.com).

The diagram in Figure 1 is merely the iconic structure of the underlying system of equations that generate behaviour during a simulation run.

This ‘gap closing method’ follows Soper (1973). Becker (1976) argued for using a ratio of post- and pre-test scores to represent relative improvement, but Soper (1976) and Highsmith (1976) disagreed.

Meta-analysis originated with Glass (1976, 1977). Before reading the seminal works, however, see Glass (2000) for an entertaining yet useful background story. Stanley (2001) documents the contribution that meta-regression analysis is making to empirical research in economics. He also provides a summary of widespread applications in other research fields where, instead of estimating regression parameters, the research goal is often the evaluation of various studies of instructional methods.

For example, the meta-analyses by Stanley (1998), Card and Krueger (1995), and Hedges et al. (1994) applied to 28, 15, and 38 studies, respectively.

Raw data and meta-analysis statistics are available upon request.