Testing Dynamic Decision Making Under Real-Time Pressure: A Scuba Diving Simulator[1]

Yaman Barlas
Industrial Engineering Department, Boğaziçi University,
Bebek, 34342, Istanbul, Turkey
Tel: +90-212-3597072 Fax: +90-212-2651800
ybarlas@boun.edu.tr

Evrim Dalkıran
Grado Department of Industrial and Systems Engineering
Virginia Tech
Blacksburg, VA 24061 USA
dalkiran@vt.edu

Abstract

Experimental dynamic decision making studies are typically carried out in environments where subjects have plentiful time before making decisions. In this research, a scuba diving simulator is developed for experimental analysis of decision making under real-time pressure, in dynamic feedback environment. In our clock-driven scuba diving simulator, subjects make decisions in real-time, continuously, which enables us to study effect of game speed (time pressure) on performance and on learning. Results show that game speed has significant effect on subjects' performances. Material and information delays are further incorporated to evaluate effects of delays on performance and learning. Both information and material delays are found significantly influential on performance. However, performance differences between delay and no-delay games decrease with practice. Since games attempt to simulate experiential learning, subjects having real diving experience may be expected to perform better than inexperienced ones. Interestingly, no statistically significant difference is found between those with scuba-diving experience and those without. An interesting feature of the game is the fact that the control problem that subjects face is under strong influence of a positive feedback loop. Combined with delays and nonlinearity, the game illustrates how complex the dynamic control problem can become even for a small, three-stock model. Performances of subjects in most trials are unstable and strongly oscillatory.

---

[1] Supported by Bogazici University Research Grant no. 02R102.

1. Introduction

Studies show that delays, feedbacks and non-linearities greatly complicate the decision making and management in dynamic feedback environments. As a result, problematic behaviors are observed due to misperceptions of feedback in complex systems. (Sterman, 1989; Brehmer, 1995; Moxnes and Saysel, 2004; Barlas and Özevin 2004).

In this study, the goal is to evaluate the effects of delays, time pressure and prior subject experience on performance, in a non-linear real-time dynamic decision making environment. A system dynamics model of scuba diving is developed and converted to an interactive game (ScubaSim) in which subjects try to stabilize at a desired depth by making air inflating/deflating decisions continuously throughout the simulation.

People make decisions using feedbacks that they receive from environment. In the existence of delays, the effect of an action is observed after a delay, so feedback clarity deteriorates. In ScubaSim, depth variable (stock) is controlled indirectly via another stock (air in jacket), constituting subtle, implicit/indirect delays. (Other explicit material and information delays are also added later to the model for explicit testing of delay effects). Most of the experimental decision making research involves gaming decisions under unlimited time (Diehl and Sterman, 1995; Aybat et al., 2004). However, decision making in reality typically involves time limits and pressures. In real life decision making, time pressure adds another complexity that may deteriorate performance (Größler,1999; Gonzalez, 2004). ScubaSim is modeled as a clock-driven simulator in which decisions are made continuously, in real-time during simulation. Another interesting feature of ScubaSim is the fact that the control problem that subjects face is under strong influence of a positive feedback loop, causing the system to be relatively unstable, hence rather hard to control –as will be discussed below.

2. Model Description

In scuba diving, diver regulates buoyancy by deflating air from or inflating air into the jacket (buoyancy compensator-BC). With a neutral buoyancy, divers stay where they are. With positive (inflated) buoyancy, diver rises without any effort whereas with negative (deflated) buoyancy diver goes deeper. Thus, controlling buoyancy is at the heart of the system. The goal is to dive and stabilize at 10 meters.

2.1. Causal-loop Diagram

An air (volume) increase in BC causes an increase in lifting force (Figure 2.1). As lifting force increases, net downward force and acceleration decrease. This results in decrease (or in less increase) in downward velocity. As velocity decreases, depth decreases and pressure decreases which eventually increases air volume. This positive feedback loop is a lifting reinforcing loop. As air is increased (decreased), the reinforcing loop starts to work and results in more increase (decrease) in volume of air in jacket. Another important component of the system is drag force. Drag force is typically formulated as a function of the velocity and shape of the body (Munson *et al.*, 2002). For our purposes, it was sufficient to formulate the drag force as a function of the velocity only. Drag force acts against to the motion. An increase in the velocity causes an increase in the magnitude of the drag force, which means a decrease in the drag force vector (more negative), which in turn decreases the net force in the direction of velocity. A net force decrease results in decreased acceleration and velocity. This relation closes the negative feedback loop, called the drag force balancing loop. Although there are some omitted other auxiliary variables and constants, fundamental variables are shown in the causal loop diagram (Figure 2.1).

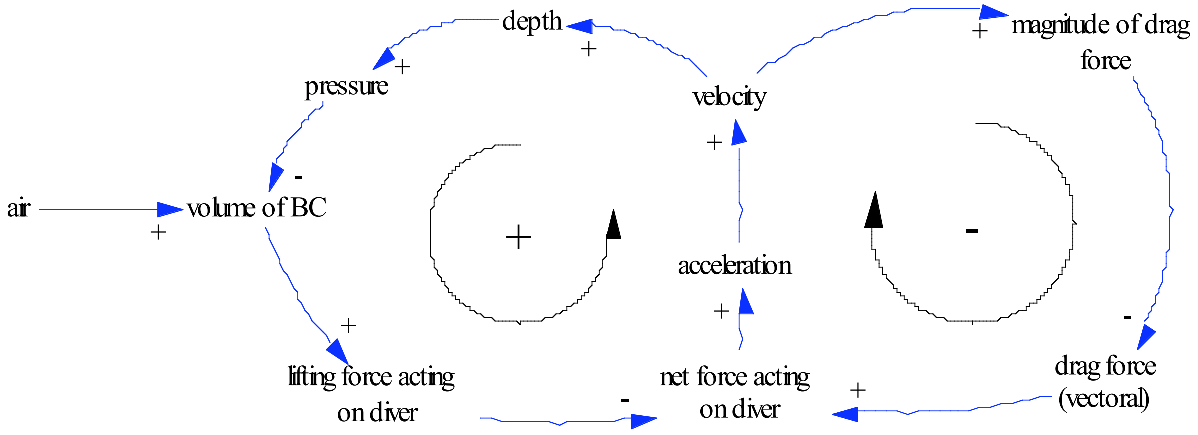


Figure 2.1. Simplified causal loop diagram of ScubaSim model

2.2. Stock-flow Diagram

The three stock variables in ScubaSim are depth, velocity and air (Figure 2.2). The rate of change of depth is velocity, which is a bi-flow. The rate of change of velocity is acceleration, also a bi-flow. The stock variable air is changed by air flow and it is a function of air adjustment decision. Disturbance outflow is used only once at t=5, to disturb the system from its equilibrium for games initialized at the desired 10 meter depth. (For games initialized at disequilibria, no disturbance is needed). Formulations of some important variables are given below. (See Appendix A for full list of equations).

The flow of air in or out of jacket as a result of player decision is given by:

$$air\_flow=Air\_Adjustment\_Decision*normal\_flow*pressure/R*T \quad \text{(moles/sec)}$$

where

Air Adjustment Decision is 0, 1 or -1 as entered by the player (or by the decision heuristic),
normal flow is 1 or 3 lt/sec depending on the flow volume level (high/low) of the game,
R is universal gas constant,
T is temperature.

Pressure is given by:

$$pressure = 1+effective\_depth/10 \quad \text{(atm)}$$

where pressure is 1 atm at the surface and increases by 1 atm for every 10 mt depth increase.

The volume of air in jacket is formulated using Ideal Gas Law:

$$volume\_lt = air*R*T/pressure$$ (liters) where *air* is the amount of air in moles

The volume_in_water is:

$$volume\_in\_water = volume\_BC\_in\_water+volume\_of\_diver\_in\_water \quad (\text{m}^3)$$

where volumes under water are calculated as a function of depth.

The lifting force is formulated by:

$$lifting\_force = -volume\_in\_water*gravitational\_constant*density\_of\_water \quad \text{(Newton)}$$

Drag force formulation is:

$$drag\_force = 1/2 * C_d * A * d * v^2 = 27.2*velocity^{\wedge}2 \quad \text{(Newton)}$$

where

v is relative velocity of the object to the fluid,

A is the cross sectional area of the object,

d is density of the fluid,

Cd drag coefficient, depends on the velocity of the object, viscosity of the fluid, shape of the body and roughness of the surface of object and is estimated as 0.8 by simulated experiments (See Dalkiran, 2006).

For the purpose of the study, A is fixed and $1/2 * C_d * A * d = 27.2$ is obtained. The direction of the drag force is opposite to motion, formulated by:

drag_force_vector = IF(velocity=0) THEN(0)
ELSE(-velocity/ABS(velocity)*drag_force) (Newton)

Net force is simply the summation of forces acting on diver and formulated by:

net_force = weight+lifting_force+drag_force_vector (Newton)

The rate of change of velocity, acceleration is formulated by Newton's Second Law:

acceleration = net_force/mass (meter/sec^2)

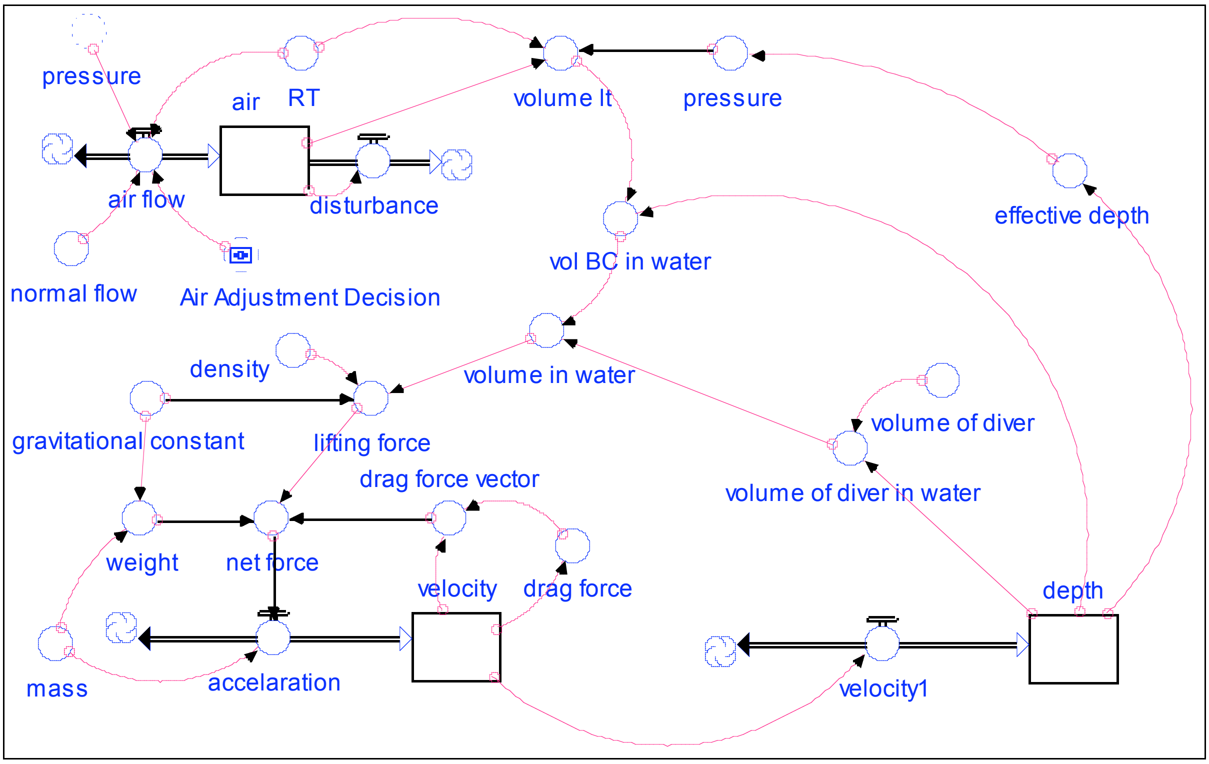


Figure 2.2. Stock-flow diagram of ScubaSim

2.3. Base Behavior of the Model

The model is tested with a wide range of different time step dt values and also with three alternative decision heuristics and real decision inputs. As a result of these verification and structure validity tests, dt=0.0125 is found appropriate with Runge-Kutta 4 integration method (See Dalkiran 2006 for extensive verification with other decision rules and dt values). Since dt affects the speed of the simulation and dt=0.01 gives better time windows to analyze time pressure, dt=0.01 is used in experiments.

It is hypothesized that dynamics of the diving can be oscillatory and one of the objectives of this study is to analyze the reasons of this problematic behavior. The model is simulated with an anchor-and-adjust heuristic in which depth(d) data is compared with desired depth(d*) value and the discrepancy is divided by depth adjustment time t_d to obtain desired velocity, $(d^* - d)/t_d$ (see Figure 2.3). This desired velocity value is compared with velocity(v) data and discrepancy is divided by the velocity adjustment time t_v which results in desired acceleration $\left(\frac{d^* - d}{t_d} - v\right)\Big/ t_v$. With the given acceleration, desired volume of air is found as

$$V^* = \frac{mass * \left(\frac{\left(\frac{d^* - d}{t_d} - v \right)}{t_v} \right) - mass * g - F_{drag}}{density * g}$$

If desired volume is larger (smaller) than actual volume, then inflating (deflating) decision is made. If desired value equals to actual volume, then neither inflating nor deflating decision is made.

$$AirAdjustmentDesicion = \begin{cases} 1 & V^* > V \\ -1 & V^* < V \\ 0 & V^* = V \end{cases}$$

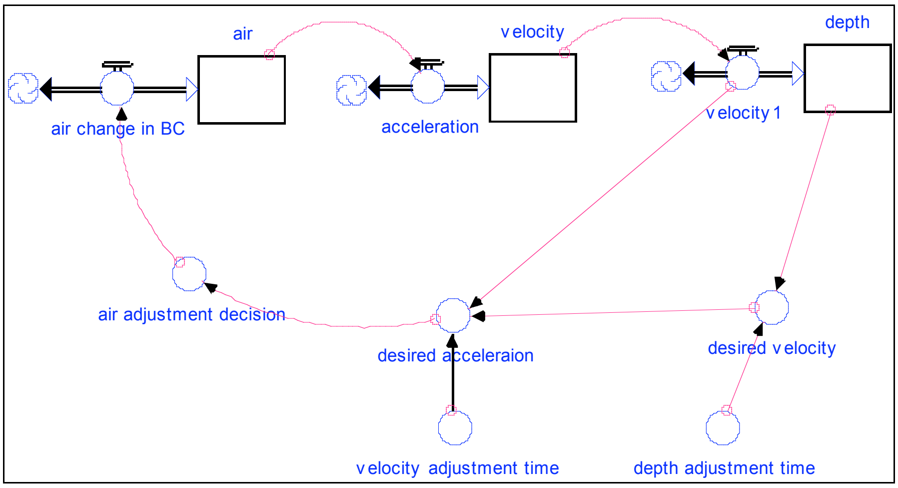


Figure 2.3. Decision heuristic structure without delay

The model is run with 1 sec, 5 sec and 25 sec depth adjustment and velocity adjustment times. The model is also run with material and information delays added to base structure. Due to space limits we are able to show only the runs with base structure, using the above anchor and adjust decision heuristic. As seen in Figure 2.4, the ratio between depth and velocity adjustment times is influential on the behavior of the model. As the ratio increases, the behavior turns to goal seeking from damping oscillations. As material delay is added and delay length is increased, amplitude of oscillations increases. With information delay added, growing oscillations may also be experienced. (Dalkiran, 2006).

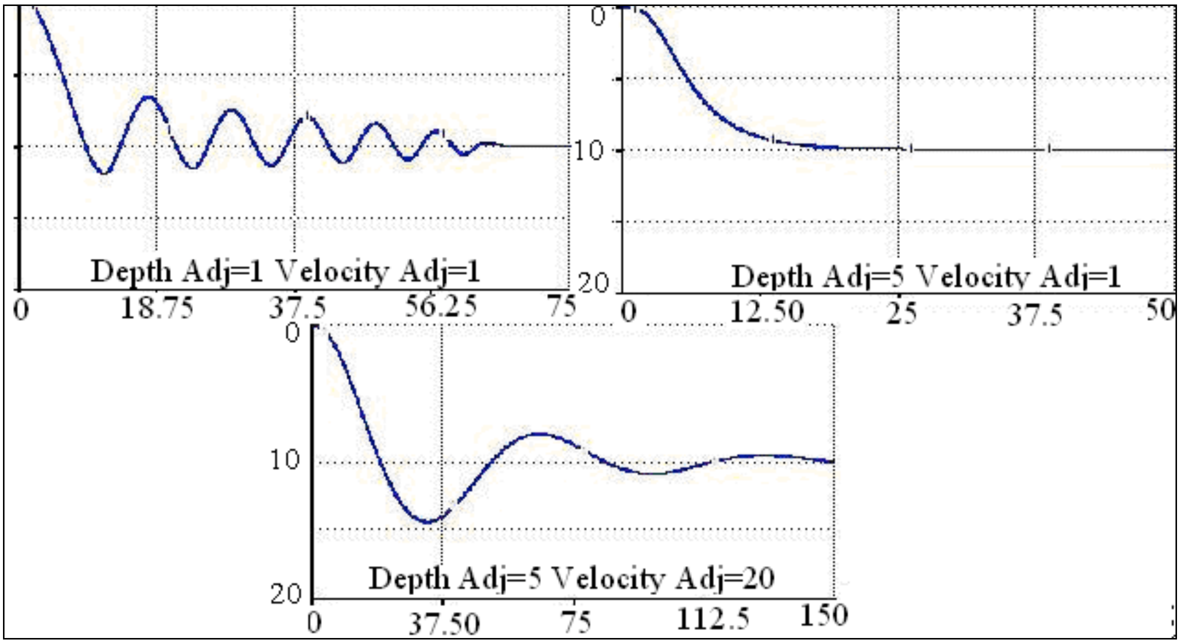


Figure 2.4. Simulation results with different depth and velocity adjustment times

3. Interactive Simulation Game (ScubaSim)

In the game version, subjects make air adjustment decisions to stabilize at desired depth and they observe depth from depth-time graph (Figure 3.1). They are expected to perceive velocity from the slope of the depth-time graph. The decisions made by subject must be entered continuously via the slider just below the graph. Once the run button is clicked, simulation start sand will does not stop until simulation is finished at t=90 seconds, except for games with pause option.

Air adjustment decision is modeled as discrete; -1 (deflate), 0 (neither deflate nor inflate) and +1 (inflate). In the left part of the slider, it is (-1), in the right part it is (1) whereas in the middle parts it is 0. The game version was once again tested and verified for robustness with alternative dt values, under different typical input decision values (see Dalkiran 2006).

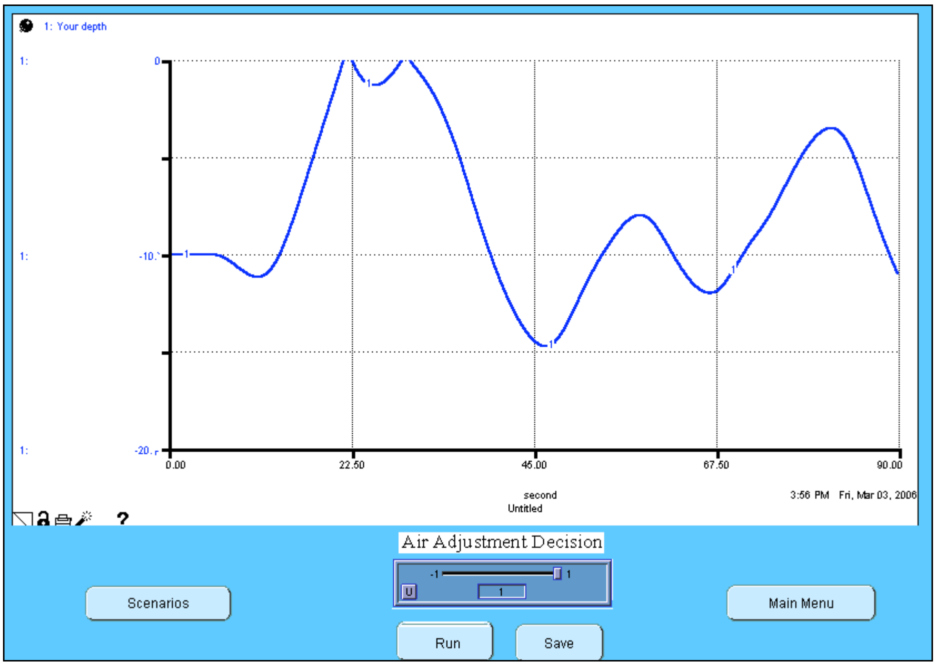


Figure 3.1. A typical example of game screen (at the end of game)

3.1. Experimental Procedure

Before experimentation, volunteer subjects read the game briefing. Since it is not easy to control slider, subjects are given time to become experienced using slider. To eliminate mechanical decision entering problems, each subject plays a pilot version of the game, not used in analysis. Subjects can also call the facilitator if they make errors due to slider and the trial is replayed from the beginning.

3.2. Effects Investigated

Effect of Game Speed

It is often hypothesized that increased game speed deteriorates performance. However experimental results do not completely support this hypothesis (Gonzalez, 2004; Größler, 1999). In ScubaSim, lengths of fast and slow game are 143 and 290 actual seconds respectively. Additionally, in one ‘extreme’ version, subjects have opportunity to “pause” the game to think about the process without any time pressure.

Effect of Air Flow Volume

In the game, subjects make inflating/deflating decision which is a binary variable. For low and high air flow volume games, 1 lit/sec and 3 lit/sec are used in simulation runs respectively. Air flow volume can be seen as control power. When the flow volume is high, the control tool is more powerful and the time needed to change the behavior of the system is relatively short compared to low air flow volume. On the other hand, with high air flow, the amount of air change may be much higher than the desired volume change and it may be too late when undesirable destabilizing effects are observed.

Effects of Information and Material Delays

The effects of material and information delays are widely studied (Aybat *et al.*, 2004; Brehmer, 1995; Diehl and Sterman, 1995; Sterman 1989). Studies show that subjects’ performances become worse when there is delay - information or material. For instance, people do not fully utilize supply line (previously ordered but not yet received stock) information when giving new orders. This behavior results in oscillations and instability. Using ScubaSim, it is possible to evaluate the effects of material and information delay in a real-time dynamic decision environment. In diving process, there may be delays in perceiving depth. To model this, a first order exponential information delay structure is used with 0.5 seconds delay length (See Figure 3.2). Additionally, there may be material delays in the flow of air through tubes. This delay is modeled as a first order exponential material delay with 0.5 seconds delay length (See Figure 3.3).

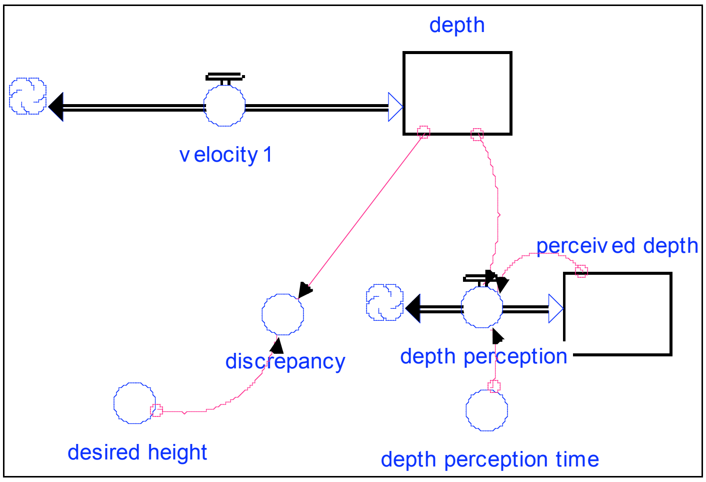


Figure 3.2. Information delay structure used in ScubaSim

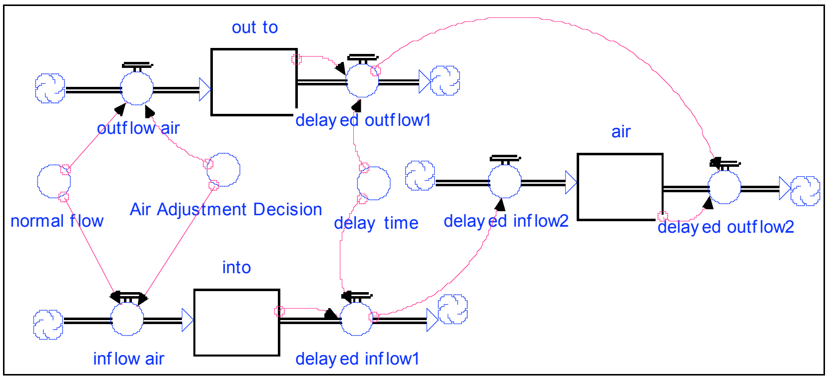


Figure 3.3. Material delay structure used in ScubaSim

The existence and length information of delays are provided to subjects at the beginning of each trial. Subjects may form an appropriate strategy to cope with delays based on this information.

Effect of Scuba Diving Experience

Subjects with diving experience probably have mental models that have been formed previously. However, their decision processes are somewhat intuitive in the sense that they may decide inflating/deflating decision by *feeling* the pressure and movement of water. It is not possible to simulate subjects' feeling the flow of water in ScubaSim. Subjects having scuba diving experience may thus perform even worse in ScubaSim due to the differences between their realities and the model used in the game.

3.3. Experimental Design

Two different experimental designs are used to analyze effects. In the first one, *Latin Square* design, 16 subjects play eight different types of the game. The required number of subjects is not very high, but the drawback of Latin Square design is that it can not analyze any interaction effects. With the results obtained from Latin Square, a second experimental design is used for further analysis. In *repeated measures* experimental design, each version of game is played by a number of subjects and each subject plays the same game six times. With this design, learning effect, effects of scuba diving experience and delays are further analyzed.

3.4. Performance Measures

To evaluate the deviation from desired depth, area between depth trajectory and 10 meters line is calculated (Dev-10-mt). If the subject is not stable at desired depth then the value of this performance measure will be high. The formulation of dev-10-mt is

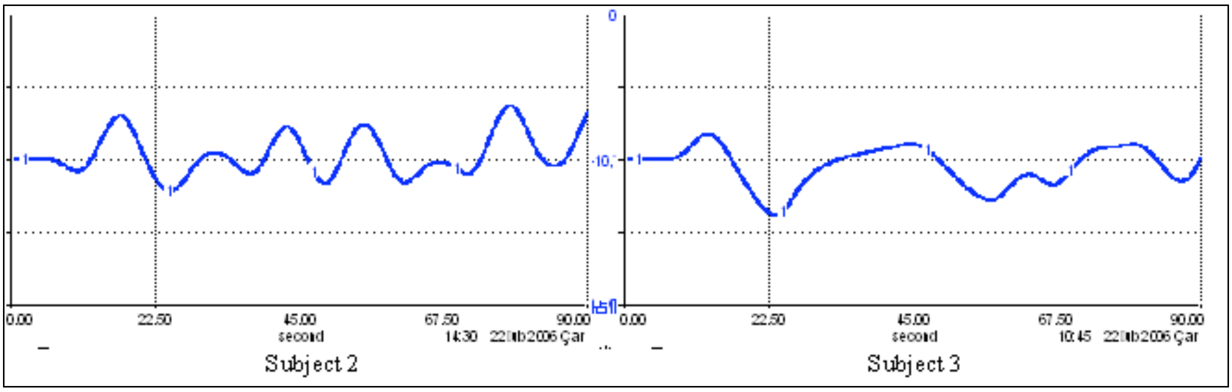
$$Dev\text{-}10\text{-}mt\ (t+dt)=Dev\text{-}10\text{-}mt\ (t)+abs[depth(t+dt)-10]*dt$$

Dev-10-mt is used as the main, default performance measure in all analysis, unless otherwise stated. Additionally, a symbolic reward is given to subjects, depending on their *dev-10-mt* scores.

The second main performance measure is the summation of amplitudes of successive cycles. The performance measure is called total amplitude of fluctuations (*amp-of-fluct*). *Amp-of-fluct* performance measure is independent of the desired depth. The formulation of *amp-of-fluct* is as follows:

$$Amp\text{-}of\text{-}fluct(t+dt)=amp\text{-}of\text{-}fluct\ (t)+abs[depth(t+dt)-depth(t)]*dt$$

Maximum deviation from 10 meters (*max-dev-10*), ±2 settling time and *deviation from 10 meters in first oscillation* are other primary performance measures. Six more performance measures are utilized in this study. However, results with respect to these measures will not be presented unless they contradict the above main measures.



Subject 2

Subject 3

Figure 3.4 Output behaviors obtained by subject 2 and 3 in game type 4

Two typical game results are shown to explain differences between *dev-10-mt* and *amp-of-fluct* performance measures. The output behaviors obtained by subject 2 and 3 in game type 4 (low speed, delay and high flow volume) can be seen in Figure 3.4. They have scores 103 and 102 respectively in *dev-10-mt*; however they have 46 and 26 in *amp-of-fluct* measure. Thus, *dev-10-mt* performance measure can not differentiate these two games whereas *amp-of-fluct* can.

4. Latin Square Experiments

In the explorative phase, Latin Square (LS) design is chosen since it can handle more nuisance factors with relatively low number of trials. The primary trial types of the game depend on the speed, level of delay, and air flow rate. The nuisance factors are subjects and experience through the trials. One drawback of the design is that interaction effect between factors can not be evaluated. However, this is only an explorative study and our main interest is to evaluate the primary factor (game type) only.

Table 4.1. Properties of games used in Latin square experimental analysis

Game type	Game speed	Delay	Flow volume
1	low	not present	low
2	low	not present	high
3	low	present	low
4	low	present	high
5	high	not present	low
6	high	not present	high
7	high	present	low
8	high	present	high

LS experiments are conducted within two different groups, depending on the initial position of the diver - 10 meters or surface. Eight subjects in each group played the eight different types of ScubaSim (Table 4.1) but with a different sequence (Table 4.2). The term (SD) denotes system dynamics knowledge. The random table is obtained with the use of program random2.exe (Byers, 1993).

Table 4.2 Order of games played by subjects (Numbers are game type numbers of Table 4.1)

	Subject1	Subject2	Subject3 (Diver)	Subject4 (Diver)	Subject5 (SD)	Subject6 (SD)	Subject7 (Diver-SD)	Subject8 (Diver-SD)
Trial 1	7	8	1	2	5	4	3	6
Trial 2	3	4	5	6	1	8	7	2
Trial 3	5	6	7	8	3	2	1	4
Trial 4	2	3	4	5	8	7	6	1
Trial 5	1	2	3	4	7	6	5	8
Trial 6	6	7	8	1	4	3	2	5
Trial 7	8	1	2	3	6	5	4	7
Trial 8	4	5	6	7	2	1	8	3

Our statistical model is as in the Equation below, where x_{ijk} is the score obtained in i^{th} trial by subject k in game type j and μ is overall mean, α_i is the effect of trial i, τ_j is the effect of game type j and β_k is the effect of subject k (Montgomery, 1997).

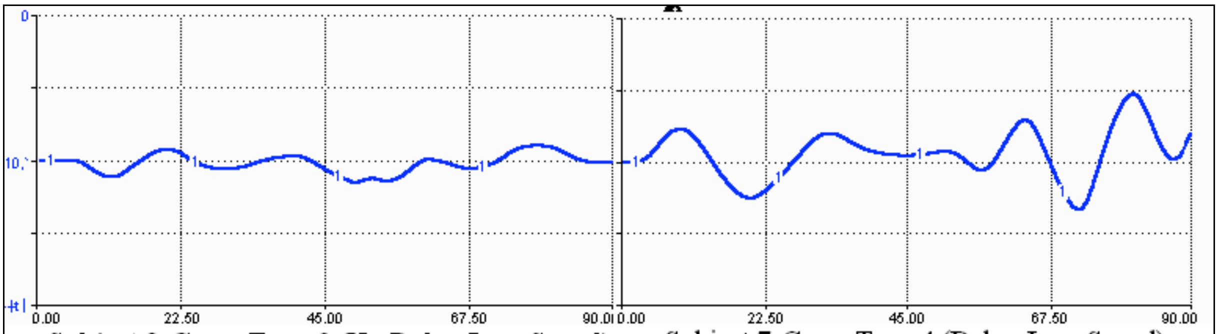
$$x_{ijk}=\mu+\alpha_i+\tau_j+\beta_k+\varepsilon_{ijk} \qquad i, j, k = 1,2...,8$$

Only two of three subscripts are required to uniquely represent an observation. For instance, when i=1 and j=1, k will automatically be 3, since game type 1 is played in first trial only by subject 3. Thus, there will be 8x8 experiments instead of 8x8x8 experiments in Latin Square design. Three main hypotheses to be tested are as follows:

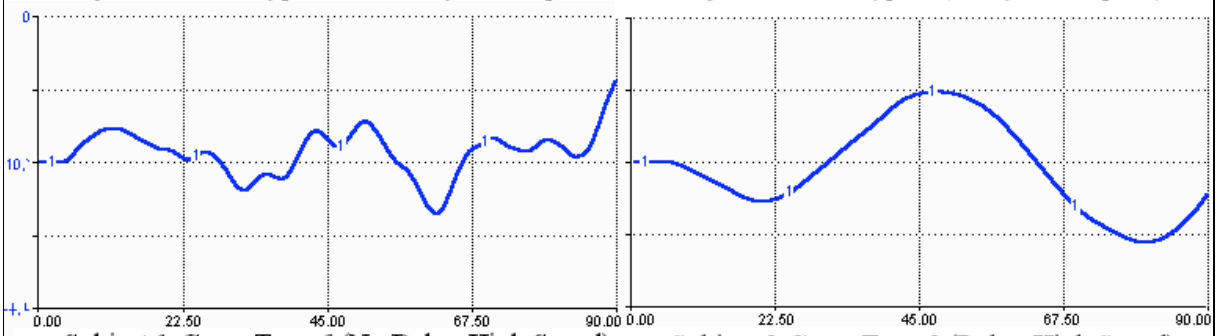
$H_0(1)$: $\alpha_1=\alpha_2=\ldots=\alpha_8$
$H_1(1)$: At least one α_i is different from others
$H_0(2)$: $\tau_1=\tau_2=\ldots=\tau_8$
$H_1(2)$: At least one τ_i is different from others
$H_0(3)$: $\beta_1=\beta_2=\ldots=\beta_8$
$H_1(3)$: At least one β_i is different from others

4.1. Games initialized at 10 meters

Before presenting statistical results, some typical output behaviors are shown to give an idea about output behaviors. At this point, it must be mentioned that comparison of these games is not unbiased such that there might also be subject differences, learning and air flow rate effects.



Subject 3-Game Type 2 (No Delay-Low Speed) Subject 7-Game Type 4 (Delay-Low Speed)



Subject 1-Game Type 6 (No Delay-High Speed) Subject 5-Game Type 5 (Delay-High Speed)

Figure 4.1. Typical results obtained in games initialized at 10 meters

When we compare the upper two games, it is seen that existence of delays deteriorates controllability of the system. Similar conclusion can be derived in the comparison of the left two games. However, in high speed game, frequency of the cycles is increased. In these three games, stability is achieved in some phases of the game. But in the high speed game with delays stability is not observed throughout the game. Although game length is not enough to reach strong conclusions, oscillation with large amplitude shows that delay and time pressure significantly increase complexity of the system.

At the first step of statistical analysis, independence and normality assumptions of residuals are checked. Due to violation of assumptions, analysis is redone with log transformation of *dev-10-mt* score (Table 4.3). Game type and practice effects and differences between subjects are statistically significant. In particular, game speed and delay are significant with respect to both measures, but flow volume and interactions are not significant in *dev-10-mt* measure. The results are almost same with other performance measures with one exception that game speed is not significant in *max-dev-10* measure.

Table 4.3. ANOVA table for games initialized at 10 meters depth (after log transformation of *dev-10-mt* scores)

		Deviation Area from 10 meters		Total Amplitude of Fluctuations		Maximum Deviation from 10 meters		±2 settling time	
	Source of Variation	SS	P-value	SS	P-value	SS	P-value	SS	P-value
GameType-Total		8.92	0	5754	0,00	135	0,00	27008	0,00
Game Type	Game Speed	1.12	0.02	1055	0,00	5	0,26	2407	0,04
	Delay	6.48	0	2603	0,00	90	0,00	21324	0,00
	Flow Volume	0.31	0.21	869	0,00	1	0,58	6	0,91
	Game Speed x Delay	0	0.91	0	0,96	1	0,64	883	0,21
	Game Speed x Flow Volume	0.35	0.19	674	0,00	12	0,08	2298	0,04
	Delay x Flow Volume	0.13	0.43	136	0,18	4	0,30	58	0,74
	Game Speed x Delay x Flow Volume	0.53	0.11	417	0,02	22	0,02	32	0,81
Subject		8.12	0	3070	0,00	137	0,00	10271	0,02
Practice		4.26	0.01	1778	0,01	92	0,00	8104	0,06
Error		8.29		3133		155		22449	
Total		29.6		13735		519		67832	

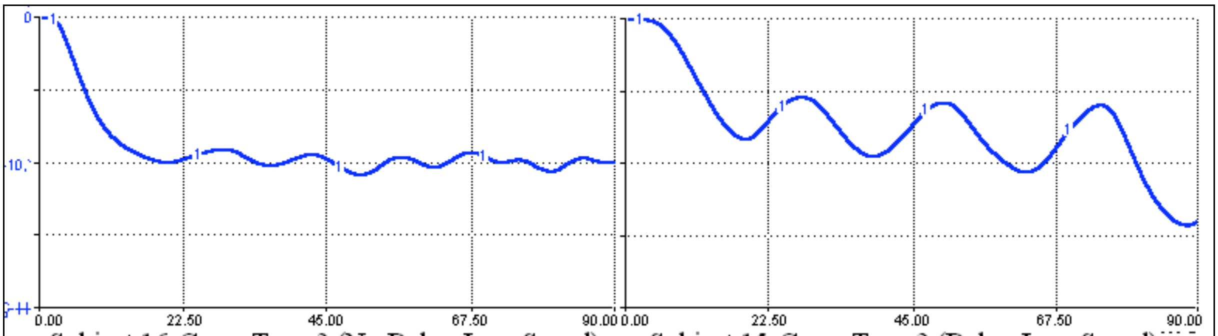
In the next analysis, effect of prior experience is studied. Subjects are divided into four groups depending on prior diving experience and system dynamics knowledge. Natural logarithms of total scores with respect to *dev-10-mt* performance measure are utilized in 2^2 factorial design with two replicates (Table 4.4). Scuba diving experience is found significant whereas system dynamics knowledge is insignificant.

Table 4.4. ANOVA table for differences between subjects: *dev-10-mt* measure (with log transformed data)

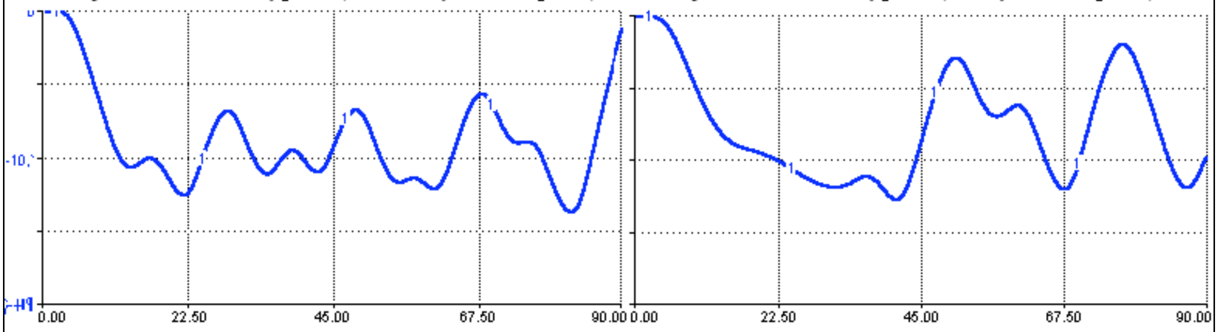
Source	SS	DF	Mean Square	F Value	Prob > F
Model	1.03	3	0.34	6.96	0.05
Scuba Diving Experience	0.93	1	0.93	18.69	0.01
System Dynamics Knowledge	0.1	1	0.1	2.1	0.22
Scuba x System Dyn.	0	1	0	0.09	0.78
Residual	0.2	4	0.05		
Cor Total	1.23	7			

4.2. Games Initialized at Surface

Although the underlying dynamics are exactly the same in surface and 10 meters games, it is worth to study the decision making when the task is driving stock from one equilibrium point to another one, instead of simply restoring it after a disturbance. As the typical game outputs are analyzed it is seen that delays and time pressure deteriorate stability and control just as in 10 meters games. However, surface games are harder to control because of the initial downwards movement after which subjects reach 10 meters depth with relatively high velocity (Figure 4.2). Subjects are able to stabilize the depth better and more often in games initialized at 10 meters, compared to the ones initialized at surface.



Subject 16-Game Type 2 (No Delay-Low Speed) Subject 15-Game Type 3 (Delay-Low Speed)



Subject 10-Game Type 6 (No Delay-High Speed) Subject 11-Game Type 8 (Delay-High Speed)

Figure 4.2. Typical results obtained in games initialized at surface

Statistical analysis shows that game speed, delay and flow volume together with subject effect are significant (Table 4.5). However, practice is found to be insignificant. As mentioned above, this may be due to the difficulty of the surface games. Learning or developing a more efficient heuristic may not be as easy as in the 10 meters games.

Table 4.5. ANOVA table for games initialized at surface

		Time Average of Deviation Area from 10 meters		Time Average of Total Amplitude of Fluctuation		Maximum Deviation from 10 meters		Deviation from 10 meter in first oscillation	
	Source of Variation	SS	P Value	SS	P Value	SS	P Value	SS	P Value
Game Type- Total		22,77	0,00	0,67	0,00	137,78	0,00	23,66	0,69
Game Type	Game Speed	1,40	0,06	0,11	0,00	32,85	0,01	0,28	0,82
	Delay	15,72	0,00	0,47	0,00	64,72	0,00	1,84	0,55
	Flow Volume	4,04	0,00	0,06	0,03	0,61	0,72	0,00	0,98
	Game Speed x Delay	0,18	0,49	0,01	0,48	2,83	0,44	4,99	0,32
	Game Speed x Flow Volume	0,49	0,25	0,00	0,94	6,52	0,25	1,46	0,59
	Delay x Flow Volume	0,74	0,16	0,00	0,92	19,45	0,05	5,00	0,32
	Game Speed x Delay x Flow Volume	0,20	0,47	0,01	0,29	10,79	0,14	10,08	0,16
Subject		5,81	0,05	1,19	0,00	76,50	0,04	52,95	0,19
Practice		1,96	0,62	0,14	0,13	42,48	0,28	17,00	0,84
Error		15,46		0,50		198,30		211,14	
Total		46,00		2,49		455,05		304,75	

4.3. Conclusions from Latin Square Experiments

The experiments with 16 subjects reveal that game speed and delays are significantly influential on the performance. Further experiments are conducted to understand individual and interaction effects of these two main factors.

The effect of flow volume is fairly different than the effects of game speed and delay. Existence of delays and high game speed deteriorate performance in all measures. However, it is seen that high flow rate improves performance if *dev-10-mt* is used and deteriorates performance if *amp-of-fluct* is used as measure. A conceptual question arises as to whether frequent oscillations with small deviations or few oscillations with large deviations are better.

In most of the experiments differences between subjects are found significant, which means individual differences bring additional variation to experimental results. Each subject is modeled as a separate block to eliminate these variations and have more accurate statistical results.

Practice effect is found significant in games initialized at 10 meters, but not in surface games. It is hypothesized that difficult control structure of surface games prevents perception, learning and performance improvement with practice.

5. Repeated Measures Experiments

The significance of time pressure, delays and practice effects in explorative study suggests further analysis of these effects. Depending on the existence of information and material delays, four different types of games; *no delay, material delay, information delay* and *both-delays*, are modeled to study delays in a broader sense. Material and information delays are modeled as continuous exponential delays with a length of 0.5 seconds. The performance deterioration due to time pressure is tested by "pause" games in which subjects are able to pause the game and use as much time as they want. Since the time pressure is highest in *both-delays* game, pause option is added to *both-delays* game structure. The analysis of the effect of prior knowledge is tested by comparing performances of scuba divers and non-scuba divers in *no delay* game. Since practice effect is found significant in explorative study, the behavior of performance throughout successive trials is further analyzed. To test the above-mentioned effects, repeated measures design is utilized which takes into account the variability due to subjects. Each game type is played by 6 to 8 subjects, and each subject played the same game 6 times.

For repeated measures analysis, independence, multivariate normality (equality of variances between groups) assumptions should hold. Although the violation of independence assumption would create serious problems, a violation of multivariate normality is not so severe. Since subjects are randomly selected and assigned to each group, independence assumption holds for the analysis. Even if there is a significant violation of multivariate normality, no additional action is taken since ANOVA is robust against deviations from normality. (Stevens, 1996).

5.1. Material and Information Delays

Brehmer (1995) examines the differences between what he calls 'dead time delay' (similar to material delay) and information delay. It is observed that information delay group performs better than the dead time delay group but worse than no delay group. Additionally, subjects' performances in information delay games become as good as the no delay group with practice. Although Yasarcan (2003), Yasarcan & Barlas (2005) prove that information delay and material delay structures are mathematically the same, perception and handling of these delays may be different.

In games with information delay, subjects observe their depth after some delay. Since they also perceive their velocity from depth trajectory, there is an implicit information delay in velocity perception. To test the effects of material and information delay and differences between delay types, two experimental designs are utilized. Firstly, no delay, material delay and information delay cases are modeled as three levels of delay factor. This design enables us to compare material and information delay. In the second design, each delay type is handled as a separate factor with two levels; delay exists or does not exist. With the second design, interaction effect can be tested.

Table 5.1. ANOVA Table for practice and delay type effects

	Dev-10-mt		Amp-of-fluct		Maximum Deviation from 10 meters		±2 Settling Time	
Source	Sig.	Partial Eta Squared	Sig.	Partial Eta Squared	Sig.	Partial Eta Squared	Sig.	Partial Eta Squared
Tests of Within-Subjects Effects								
practice	0,00	0,31	0,00	0,45	0,00	0,30	0,00	0,21
practice * Delay Type	0,12	0,13	0,76	0,06	0,83	0,05	0,01	0,19
Tests of Between-Subjects Effects								
Intercept	0,00	0,94	0,00	0,94	0,00	0,93	0,00	0,98
Delay Type	0,00	0,55	0,02	0,31	0,00	0,41	0,00	0,58

Starting with the first design, it is observed that delay type and practice effect are significant but practice-delay type interaction effect is not (Table 5.1). The pattern of performance through trials is not affected by the delay type. As scores obtained in no delay, material and information delay cases are compared, it is seen that the difference between material and information delay is not statistically significant but performances in these two types of games are significantly worse than performances in no delay game (Table 5.2).

Table 5.2. Mean scores in no delay, information delay and material delay games: *dev-10-mt* and *amp-of-fluct*

	Deviation Area from 10 meters			Total Amplitude of Fluctuations		
Delay Type	Mean	95% Confidence Int.		Mean	95% Confidence Int.	
Information	198,58	173,9	223,26	45,72	38,88	52,55
Material	169,4	144,72	194,08	45,96	39,13	52,8
No	99,36	74,68	124,04	30,94	24,1	37,77

Continuing with the second design, as seen in (Figure 5.1) mean score obtained in information delay equals to mean score obtained in both-delay case in the 1st trail. This situation can not be explained logically, but the reason may be the small sample size. At the 2nd, 3rd and 4th trials, an additive effect of delays is observed. In other words, mean scores obtained in *both-delays* games are essentially equal to the material delay effect (difference between material delay and no delay) and information delay effect (difference between information delay and no delay) added to the *no delay* score (a base score). This relation does not hold in the last two trials with respect to *dev-10-mt* score.

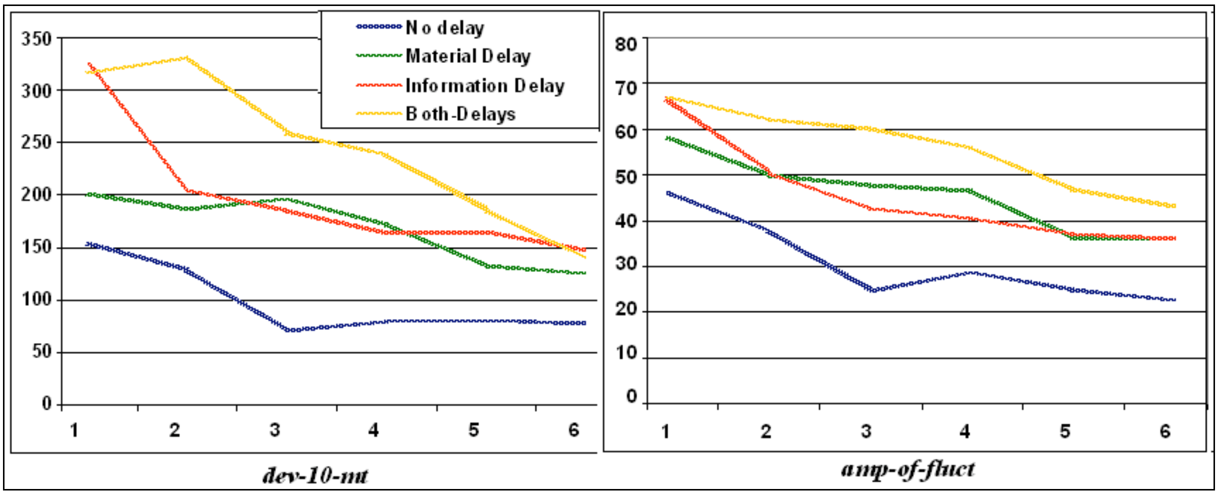


Figure 5.1: Mean scores obtained in games with respect to trials: *dev-10-mt* and *amp-of-fluct*

The results of statistical analysis show that practice, material and information delays are significantly influential on performance. However, material-information delay interaction is not significant. (Practice) x (material delay) and (practice) x (information delay) interaction effects are significant for *dev-10-mt* performance measure (Table 5.3). In the existence of information delay, performance improvements through trials are higher compared to the performance improvement in the absence of information delay. In material delay, in trial 3 and 4 performance improvement is similar, regardless of the existence or absence of material delay. But in trial 5 and 6 further improvement is observed when there is material delay, whereas performance remained nearly constant in the absence of material delay.

Table 5.3. ANOVA Table for practice, material and information delay effects

	Dev-10-mt		Amp-of-fluct		Maximum Deviation from		±2 Settling Time	
Source	Sig.	Partial Eta Squared	Sig.	Partial Eta Squared	Sig.	Partial Eta Squared	Sig.	Partial Eta Squared
Tests of Within-Subjects Effects								
practice	0,00	0,37	0,00	0,46	0,00	0,36	0,00	0,16
practice * Information	0,09	0,07	0,87	0,01	0,52	0,03	0,06	0,09
practice * Material	0,06	0,09	0,19	0,06	0,24	0,05	0,04	0,10
practice * Information * Material	0,42	0,04	0,93	0,01	0,81	0,02	0,13	0,06
Tests of Between-Subjects Effects								
Intercept	0,00	0,95	0,00	0,93	0,00	0,95	0,00	0,98
Information Delay	0,00	0,56	0,02	0,20	0,00	0,39	0,00	0,38
Material Delay	0,00	0,36	0,01	0,21	0,01	0,23	0,00	0,38
Information * Material	0,46	0,02	0,63	0,01	0,28	0,05	0,01	0,25

5.2. Scuba Diving Experience

Subjects may gain experience and learn about the dynamics of scuba diving by successive games with ScubaSim. Conversely, it can be hypothesized that subjects with real life diving experience may perform better in the game than the subjects without experience. To test the hypothesis, performances obtained by scuba-divers and non scuba-divers in *no delay* game are compared.

Table 5.4. ANOVA Table for scuba diving experience

	Deviation Area from 10 meters		Total Amplitude of Fluctuations		Maximum Deviation from 10 meters		±2 Settling Time	
Source	Sig.	Partial Eta Squared	Sig.	Partial Eta Squared	Sig.	Partial Eta Squared	Sig.	Partial Eta Squared
Tests of Within-Subjects Effects								
practice	0	0,35	0	0,44	0,03	0,19	0	0,28
practice * Scuba	0,62	0,06	0,6	0,06	0,37	0,09	0,52	0,07
Tests of Between-Subjects Effects								
Intercept	0	0,86	0	0,83	0	0,8	0	0,8
Scuba	0,32	0,09	0,65	0,02	0,34	0,08	0,49	0,05

Analysis shows that diving experience does not have significant effect on the performance of subjects (Table 5.4). Practice effect is significant for all performance measures whereas (practice) x (scuba) interaction effect is not significant. Performance statistically differs among trials but the pattern of performance is independent of scuba diving experience. Thus, all subjects have similar learning patterns (Figure 5.2).

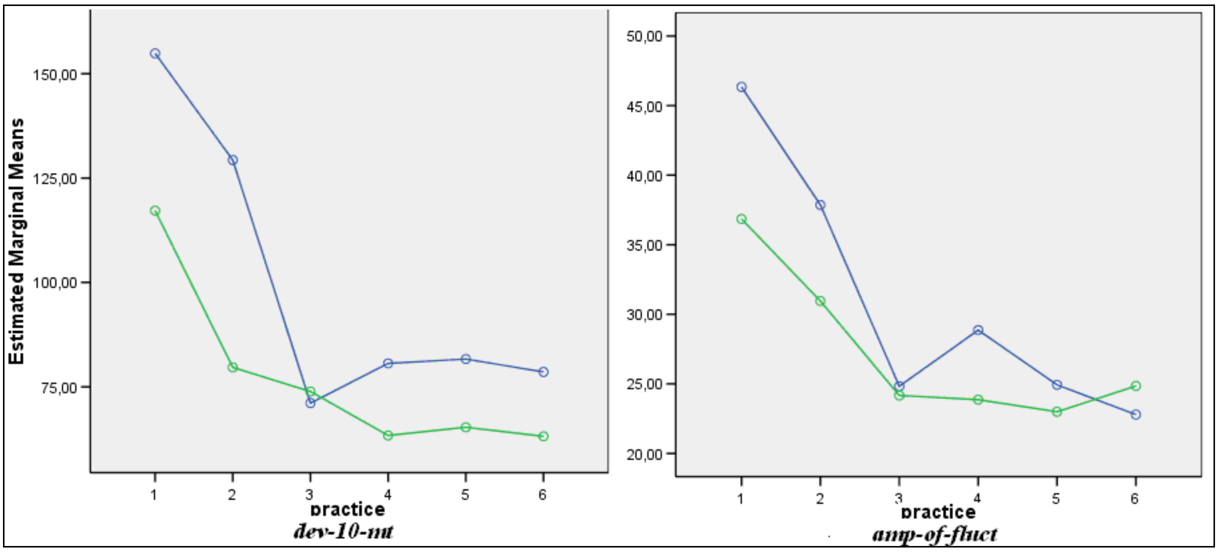


Figure 5.2. Mean scores of scuba divers and non scuba divers: *dev-10-mt* and *amp-of-fluct*

5.3. Effect of Pause Option

Größler(1999), Gonzalez (2004) and Brehmer (1990) designed experiments to account for the effect of time pressure or time limit on dynamic decision making. In a similar way, to test the effect of time pressure on performance, ScubaSim is played with pause option as an opportunity to pause the game indefinitely whenever subjects want additional time to think.

Subjects are explicitly informed about the pause option and encouraged to use it. However, as games with and without pause option are compared, no significant difference is observed (Table 5.5). Subject 47 used pause option only once and only in trial 2. He stated that he did not need to use pause, because his strategy would not change with additional time to think. He also added that with pause option, he lost attention. After first trial, subject 48 said that she could not decide when to use pause option. After second trial, she concluded that pause option is not necessary and she did not use pause option in any trials. Like subject 48, subject 49, 50, 51, 52 and 53 stated that they did not need to use pause option in any of the 6 trials. To conclude, none of the subjects felt using “pause” button was necessary or helpful.

As the statistical analysis is carried out, it is observed that performance difference between two groups is not significant. On the one hand, subjects may lose their sense of the system dynamics when decisions are made in discrete time points, after long pauses. But on the other hand, in many real life situations it is only possible to adjust decisions at specific time points, not continuously at all times. Our results show that first the type of the simulator (whether continuous or discrete decision interval) and then the length of the decision interval must be carefully selected in designing realistic simulators.

Table 5.5. ANOVA Table for pause effect

	Deviation Area from 10 meters		Total Amplitude of Fluctuations		Maximum Deviation from 10 meters		±2 Settling Time	
Source	Sig.	Partial Eta Squared	Sig.	Partial Eta Squared	Sig.	Partial Eta Squared	Sig.	Partial Eta Squared
Tests of Within-Subjects Effects								
practice	0.00	0.37	0.00	0.35	0.00	0.40	0.39	0.09
practice * Pause	0.43	0.08	0.38	0.09	0.23	0.11	0.20	0.12
Tests of Between-Subjects Effects								
Intercept	0.00	0.89	0.00	0.93	0.00	0.96	0.00	1.00
Pause	0.68	0.02	0.85	0.00	0.70	0.01	0.57	0.03

6. Post Game Questionnaire

Each subject answered the questionnaire after the first and the last trial. Subjects gave the answers in Table 6.1, as reasons of difficulties they experienced. (See Appendix B). ‘Several interacting factors’ explained in the manual are chosen 42 times as the reason for unsuccessful results. It is followed by game speed, weak PC game playing skills and existence of delays. It must be mentioned that delay is not presented as an option, but many subjects state independently that existence of delays was a difficulty factor.

Table 6.1. Distribution of factors attributed to the difficulties in obtaining successful results

	10 mt		Surface		No delay		Material		Information		Both		Pause		Scuba		All		**Total**
	1st	2nd	1st	2nd	1st	2nd	1st	2nd	1st	2nd	1st	2nd	1st	2nd	1st	2nd	1st	2nd	
Complexity of slider	2	2	1	0	2	2	2	2	0	1	0	0	0	0	2	2	9	9	**18**
Speed of the game	1	4	4	2	2	1	2	1	1	1	1	1	2	1	2	2	15	13	**28**
Weak PC game skills	2	1	2	1	1	0	2	1	3	2	0	0	2	2	2	2	14	9	**23**
Several interacting factors	3	3	2	1	2	3	4	3	3	1	3	3	2	3	3	3	22	20	**42**
Hidden/unknown factors	2	1	1	1			1	1	4	3	0	0	0	0	1	1	9	7	**16**
Other-Delay	1	3	3	3			1	1		1	2	2	4	4	1	0	12	14	**26**
Other-Constant flow									1	1	1	1					2	2	**4**

Subjects also chose hidden/unknown factors 16 times, which can be related to next question: About the existence of external factors/forces like currents, winds, unstable regions in water or any other random factors, 25 of 118 responses state that there are external factors for sure, while another 25 state that there may be some external forces. We can say that almost half of the subjects perceive that there are hidden external forces and almost one fifth of the subjects state these unknown factors as a reason for their difficulties in obtaining successful results.

7. Conclusions

In this study, effects of time delays, time pressure, prior real-life experience and game practice on decision performance of subjects in a dynamic feedback environment are analyzed. A scuba diving simulator (ScubaSim) is developed for experimental analysis of dynamic decision making under real-time pressure, in non-linear feedback environment. In this clock-driven scuba diving simulator, subjects make decisions in real-time, continuously, which enables us to study effect of game speed (time pressure) on performance and on learning. In Latin Square experimental design, the negative effect of delay is found highly significant. For repeated measures experiments, delay structures are added to basic structure one by one: material delay, information delay and both-delays. It is found that material delay and information delay have significant effects on performance and that there is no significant difference between material delay and information delay. Thus, type of delay (material/information) does not significantly affect the performance. Game speed (time pressure) is another game attribute found significant in Latin square experiments. Thus, as an extreme case of slow game, pause option is also incorporated in repeated measures experiments. However, subjects did not use the pause option as they felt that they would lose sense of time continuity when pause option was used. Thus there was no significant difference between games with and without pause option.

Another significant effect on performance is practice. In repeated measures experiments, each subject played six times successively the same game. In all game types, one or two-step improvements are found statistically significant.

Both in Latin Square and Repeated Measures analyses, there are differences between subjects. Blocking is used to obtain clearer results. A strong finding about subject differences is that the effect of scuba diving experience on performance is not statistically significant.

Among the different behavior types observed in the games, oscillations take the first place. In about 70 percent of the games, oscillations are significant. This ratio drops in games without delay structures and increases with delay structures. A ratio as high as 70 percent, shows that subjects do experience a management/control problem. But there are also some exceptional subjects who avoid oscillations in delay games while others who experience oscillations in no delay games. Decision processes of the subjects should be analyzed more extensively to discover the differences between decision strategies of subjects who can avoid oscillations and those who cannot. One of the interesting future research areas would be formulating and testing the decision process of subjects in this real-time dynamic decision environment. Since decisions are given continuously in real-time, it is not straightforward to determine a decision function for subjects. It would be interesting to explore whether the heuristics applicable in discrete time and/or batch decision making without time pressure are also applicable in real-time, continuous decision making.

REFERENCES

Aybat, S., S. Daysal, B. Tan and F. Topaloglu, 2004, "Decision-making tests with different variations of the stock management game", *Proceedings of the 22nd International System Dynamics Conference*, Oxford, United Kingdom.

Barlas, Y., 1996, "Formal Aspects of Model Validity", *System Dynamics Review*, Vol. 12, No. 3, pp.183-210.

Barlas, Y., & Ozevin, M. G. 2004. Analysis of stock management gaming experiments and alternative ordering formulations. *Systems Research & Behavioral Science,* Vol. 21, Issue 4, 439-470.

Brehmer, B.,1990,. "Strategies in real time, dynamic decision making", in R.M. Hogart (Ed.), *Insights in decision making*, pp. 262-279, University of Chicago Press, Chicago, IL.

Brehmer, B., 1995, "Feedback delays in complex dynamic decision tasks", in P.A. Frencsh and J. Funke (Eds.), *Complex Problem Solving: The European Perspective,* pp.103-130, Lawrence Erlbaum Associates, Hillsdale, NJ.

Byers, A. J.,1993, "Randomization Algorithms in basic for experimental design", *Comput. Biol. Med.* 23, 2.

Busemeyer, J. R. and J. G. Johnson, 2004, Computational models of decision making. in D. Koehler & N. Harvey (Eds.) *Handbook of Judgment and Decision Making*, Ch. 7, pp 133-154, Blackwell Publishing Co.

Dalkiran, E., 2006, Scuba Diving Simulator: Testing Real-time Decision Making in a Feedback Environment., M.S. Thesis, Bogazici University.

Diehl, E., and J. D. Sterman, 1995, "Effects of feedback complexity on dynamic decision making", *Organizational Behavior and Human Decision Processes*, 62(2), 198-215.

Fluent 6.1 User's Guide, Fluent Incorporated, http://www.fluent.com

Gonzalez, C., 2004, "Learning to make decisions in dynamic environments: Effects of time constraints and cognitive abilities", *Human Factors*, 46(3), 449-460.

Gonzalez, C., 2005, "Decision support for real time, dynamic decision making tasks", *Organizational Behavior and Human decision Processes*, 96, 142-154.

Größler, A., 1997, "As Time Goes By – Self-Proceeding Management Simulations", in Y. Barlas, V G. Diker and S. Polat (Eds.), *15th International System Dynamics Conference – Systems Approach to Learning and Education into the 21st Century,* 371–374, *Istanbul*.

Größler, A., 1999, "The Influence of Decision Time on Performance in Use of a Business Simulator", in R. Y. Cavana et al (Eds.), *Systems Thinking for the Next Millennium – The Proceedings of the 17th International Conference of the System Dynamics Society,* Wellington.

Montgomery, D.C., 1997, *Design and analysis of experiments*, 4th edition, John Wiley Sons, New York

Moxnes E., and A. K. Saysel, 2004, Misperceptions of global climate change: information policies, WPSD 1/04. System Dynamics Group, University of Bergen.

Munson, B.R., T.H. Okiishi and D.F. Young, 2002, *Fundamentals of Fluid Mechanics*, 4th Edition, John Wiley Sons, New York

Sterman, J. D., 1989, "Modeling managerial behavior: Misperceptions of feedback in dynamic decision making", *Management Science*. 35(3), 321-339.

Stevens, J., 1996, *.Applied Multivariate Statistics for the Social Sciences*, Lawrence Erlbaum Associates, Mahwah, NJ

Yasarcan H. 2003. Feedback, Delays and Non-linearities in Decision Structures. PhD Thesis, Bogazici University.

Yasarcan, H., & Barlas Y. 2005. A generalized stock control formulation for stock management problems involving composite delays and secondary stocks. *System dynamics review*, *21*(1), 33–68.

APPENDIX A: MODEL EQUATIONS

air(t) = air(t - dt) + (air_flow - disturbance) * dt (moles)
INIT air = pressure*(90-volume_of_the_diver*1000)/(R*T) (moles)
INFLOWS:air_flow=Air_Adjustment_Decision*normal_flow*pressure/(R*T) (moles/sec)
OUTFLOWS: disturbance = PULSE(air/4,5,0) (moles/sec)
depth(t) = depth(t - dt) + (velocity1) * dt (meters)
INIT depth = 10 (meters)
INFLOWS:velocity1 = velocity (meter/sec)
velocity(t) = velocity(t - dt) + (acceleration) * dt (meter/sec)
INIT velocity = 0 (meter/sec)
INFLOWS: acceleration = net_force/mass (meter/sec^2)
Air_Adjustment_Decision = 0
density_of_diver = 1070 (kg/m^3)
density_of_water = 1000 (kgs/m^3)
drag_force = 27.2*velocity^2 (Newton)
drag_force_vector = IF(velocity=0)
THEN(0)
ELSE(-velocity/ABS(velocity)*drag_force) (Newton)
effective_depth = IF(depth<-0.5)
THEN(0)
ELSE(IF(depth<0)
THEN((0.5+depth)/2)
ELSE(depth+0.5/2)) (meters)
gravitational_constant = 9.81 (meter/sec^2)
lifting_force = -volume_in_water*gravitational_constant*density_of_water (Newton)
lt_m3_conversion_factor = 1000 (liters/meters^3)
mass = 90 (kgs)
net_force = weight+lifting_force+drag_force_vector (Newton)
normal_flow = 1 (liters/sec)
pressure = 1+effective_depth/10 (atm)

```
RT = 25 (atm/(liters*moles))
volume_BC_in_water = IF(depth<-0.5)
                         THEN(0)
                         ELSE(IF(depth<0)
                                 THEN(volume_m3*(0.5+depth)/0.5)
                                 ELSE(volume_m3)) ($m^3$)
volume_of_diver_in_water = IF(depth<-1.5)
                         THEN(0)
                         ELSE(IF(depth<0)
                                 THEN(volume_of_the_diver*(1.5+depth)/1.5)
                                 ELSE(volume_of_the_diver)) ($m^3$)
volume_in_water = volume_BC_in_water+volume_of_diver_in_water ($m^3$)
volume_lt = air*RT/pressure (liters)
volume_m3 = volume_lt/lt_m3_conversion_factor ($m^3$)
volume_of_the_diver = mass/density_of_diver ($m^3$)
weight = mass*gravitational_constant (Newton)
```

APPENDIX B: POST GAME QUESTIONNARE

1. Was the objective the task involved in the game clear?

2. Did you follow a specific strategy that you can describe? (Did your strategy differ in different versions of the game?)

3. If you had difficulties in obtaining successful results, do you think these difficulties are due to;
    - Complexity of mechanics of entering decisions through slider
    - Speed of the game
    - Your weak computer/video game playing skills(as in PC or Playstation games)
    - Complexity of conceptually coming up with proper decisions, due to several interacting factors/factors explained in the manual
    - Hidden/unknown difficulty factors in the game
    - Other

4. Do you think there were any external factors/forces like currents, winds, unstable regions in water or any other random factors during the game?

5. Do you find yourself successful? If not, how do you think you could improve?
6. Would any additional info increase your success in the game?
    - Velocity data
    - Air volume data
    - Force data
    - Other