

The pond-and-duckweed problem : three experiments on the misperception of exponential growth

Wagenaar, W.A.; Timmers, Han

Published in:
Acta Psychologica

DOI:
[10.1016/0001-6918\(79\)90028-3](https://doi.org/10.1016/0001-6918(79)90028-3)

Published: 01/01/1979

Document Version
Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

- A submitted manuscript is the author's version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

Citation for published version (APA):
Wagenaar, W. A., & Timmers, H. (1979). The pond-and-duckweed problem : three experiments on the misperception of exponential growth. *Acta Psychologica*, 43(3), 239-251. DOI: 10.1016/0001-6918(79)90028-3

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

THE POND-AND-DUCKWEED PROBLEM; THREE EXPERIMENTS ON THE MISPERCEPTION OF EXPONENTIAL GROWTH

Willem A. WAGENAAR*

Institute for Perception TNO, Soesterberg, The Netherlands

and

Han TIMMERS

Institute for Perception Research, Eindhoven, The Netherlands

Received July 1978

The representation of duckweed multiplying itself in a pond is used as a research paradigm to study underestimation of exponential growth. The advantage of this paradigm is that the growth process is presented in a direct non-numerical way. The results show that the underestimation observed in earlier studies where growth was presented by means of tables or graphs, occurs in the pond-and-duckweed situation as well. By manipulating the way the process is presented it is possible to obtain some insight into the sampling strategies used by the subjects when they subjectively extrapolate the perceived processes. These experiments lead to the conclusion that subjects base their extrapolations on three or four samples only.

1. Introduction

Many problems the world as a whole is faced with are related to growth. Economic growth and growing populations induce shortages of space, energy, raw materials and food, and an increase of such seemingly unrelated quantities as cost of living, pollution, crime rate, rate of divorce and number of scientific publications. Usually these processes start with an exponentially growing phase; later saturation should occur, leading to a levelling off, thus producing the so-called logistic curve (de Solla Price 1963). Since any attempt to control such processes will depend on the cooperation of individual citizens some

* Requests for reprints should be sent to Dr. Willem A. Wagenaar, Institute for Perception TNO, Kampweg 5, Soesterberg, The Netherlands.

insight in the intuitive evaluation of exponential growth might contribute to the solution of growth problems.

In experiments reported before (Wagenaar and Sagaria 1975) it was shown that exponential growth is considerably underestimated when the processes are presented by means of tables or graphs. The general finding was that people tend to extrapolate exponentially, that is with a constant multiplier for successive steps, but with an exponent that is too small. The real exponent was weighed with a factor 0.20 for tabular presentation and 0.04 for graphical presentation. Neither mathematical sophistication of the subjects nor experience with growth processes changed this effect.

One important feature of tabular and graphical representations is that they exclude the element of time: a dynamite explosion develops in microseconds; the process of milk boiling over occurs within a few seconds; colonies of influenza bacilli develop within some hours; price indices grow over years and populations accumulate over ages. Presentation of these processes by tables or graphs neglects these differences because a large part of the history of a process is presented; moreover, in reality people are often only confronted with the present state of a process (cf. high prices, full parking lots, dirty rivers) while for the history they have to rely on memory. Does this factor change the effect of underestimation?

A study on the extrapolation of processes covering several years (Wagenaar and Timmers 1978) showed a marked effect of whether or not the past history was presented. In the present experiment the past history is limited to a maximum of about five minutes.

Another aspect of tables and graphs is the quantitative nature of the representation; many processes do not present themselves in a pronounced quantitative way: one cannot directly count pollution of water, and one normally will not count the number of cars involved in a traffic jam. Therefore extrapolation of such processes is almost certainly not mediated by simple mathematical algorithms. Does this exclusion induce different strategies leading to different extrapolations? Or will the model proposed by Wagenaar and Sagaria (constant underestimation of the exponent) also be valid for non-quantitative representations?

Experiment 1

Method

Apparatus and stimuli

The basic experiment was inspired by the fable about the Chinese mandarin who as a youth planted some duckweed in a pond. The duckweed doubled every five years. When the mandarin was 70 years old he was quite satisfied to observe that one eighth of the pond was covered. He did not realize that in the 15 years to come one quarter, one half and finally all of the pond would be covered with weed. Never during the process was the mandarin worried by the growth, although he had a lifetime of observations. A similar process was simulated on the scope display of a PDP-8. The pond was a 10×10 cm square that could contain 256 small squares (duckweeds) in rows of 16. The number of duckweeds (n) increased as a function of time (t) according to:

$$n = a e^{10 b \frac{t}{d}} - a \quad (1)$$

In this formula n is an integer number, and t is time in seconds, running from 0 to d , the time necessary to get the pond filled. Values of d used in this experiment were 60, 120, 240 or 480 sec. The completely filled pond was never shown to the Ss, since the presentation was stopped at $t_{\max} = 0.33 d$, $0.50 d$ or $0.67 d$. According to formula (1) we obtain $n = 0$ at $t = 0$ sec; a was chosen such that $n = 256$ at $t = d$. The values of b were $b = 0.1$, 0.3 and 0.5 ; it should be kept in mind that these values have no absolute meaning: if we omit the coefficient 10 in the exponent we would obtain the same function by choosing $b = 1$, 3 and 5 respectively. The value 10 was chosen because this way the total time necessary to get the pond filled could be easily thought of as being subdivided into ten steps; thus comparability with previous experiments on tabular and graphical representations was assured.

A trial started by presentation of the empty large square. Then the duckweeds appeared starting left in the bottom row, filling one row after the other. The time between the appearances of duckweeds n and $n + 1$ is expressed by

$$\Delta t = \ln \left(\frac{n + a + 1}{n + a} \right) \times \frac{d}{10b} \quad (2)$$

where n = ranknumber of the first duckweed of each pair. The longest interval (43.6 sec) occurs before the appearance of the first duckweed in the condition $b = 0.5$, $d = 480$ sec. The shortest interval (0.2 sec) is between the two first duckweeds in the condition $b = 0.3$, $d = 60$ sec, $t_{\max} = 0.67 d$.

When the presentation stopped at t_{\max} the Ss had to indicate on a linear scale which proportion of the time d had elapsed. If they estimated that the additional time needed to get the pond completely filled was about equal to the time consumed so far, they would place their marks halfway on the scale. The length of the scale was 15 cm.

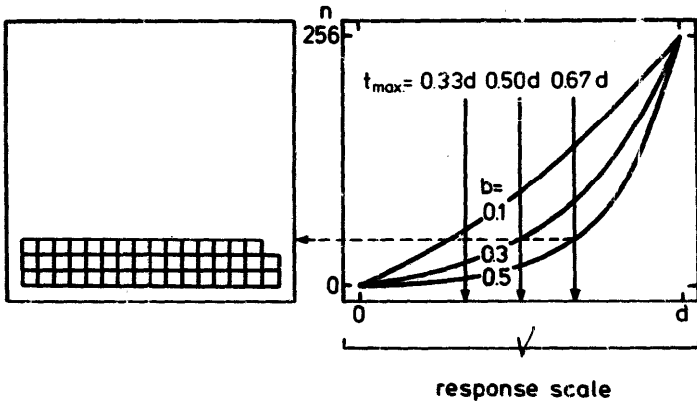


Fig. 1. The pond-and-duckweed paradigm. The number of squares has been accumulated according to the function $n = 1.74 e^{\frac{10(0.5)t}{d}} - 1.74$, until $t = 0.67d$ ($n = 47$). The question now is: how far are we between the beginning and the moment at which the pond will be completely filled? A subject who thinks that we are about halfway (mark in the middle of the response scale) is assuming a growth function with $b = 0.3$. In that case $0.3/0.5 = 60\%$ of the exponent is taken into account.

An example of the scoring method is presented in fig.1. The responses are translated into scores on the assumption that the Ss will perceive the exponential character of the process; only the value of the exponent may be misperceived. In that case a response can be translated into the exponent of the growth function as they are perceived. The quotient of this value and the veridical exponent indicates what proportion β of the exponent is taken into account. Later on it will be shown that the curvilinear growth was detected in all conditions.

Subjects

The Ss were 36 male and female students from the State University of Utrecht. They were paid Dfl. 25,- for the completion of the experiment.

Procedure

All Ss produced 36 estimates (3 values of b , 3 values of t_{\max} , 4 values of d , combined factorially). The conditions were presented according to a digram-balanced latin square (Wagenaar 1969). After presentation of the oral instruction it was checked whether the S could use the response scale correctly (e.g. if they were told that the additional time required to get the pond completely filled was one half of the time elapsed thus far, they should put a mark at two thirds of the response scale).

Results

Results are presented in figs. 2 and 3. The effect of b ($F_{2, 70} = 14.5, p < 0.01$) shows that the underestimation, which is absent at $b = 0.1$, increases rapidly with

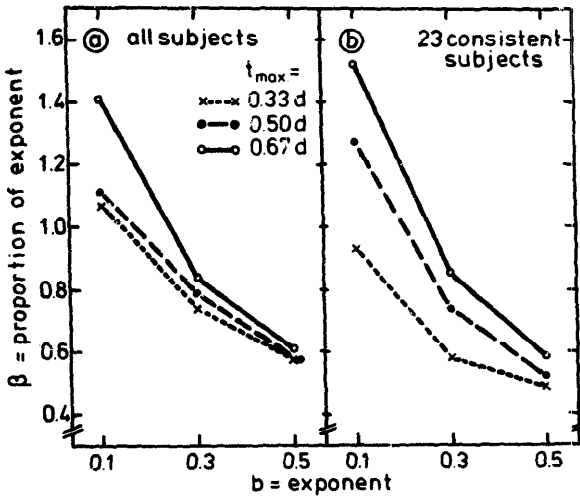


Fig. 2. Results of experiment 1; effects of t_{max} .

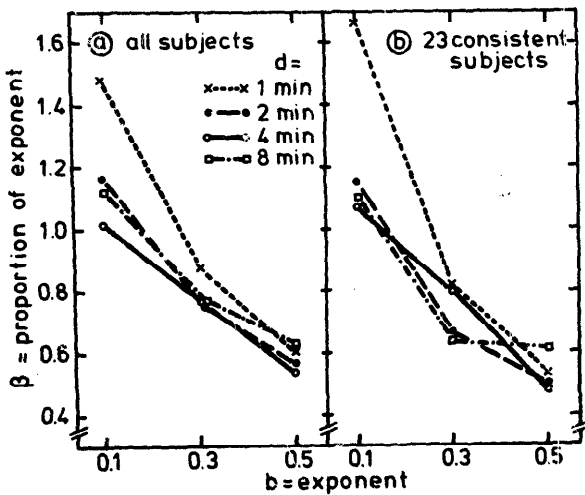


Fig. 3. Results of experiment 1; effects of d .

b ; the previously reported values of $\beta < 0.20$ when $b > 1.0$ are well in line with these results. The effect of t_{\max} was not significant ($F_{2, 70} = 1.99, p > 0.10$), but the interaction with Ss was highly significant ($F_{70, 420} = 3.29, p < 0.001$). This might mean that the effect of t_{\max} which is marginally present in fig. 2a is obscured by the results of a few Ss . After inspection of the raw responses it became obvious that 13 Ss behaved very inconsistently; that is, within sets of three conditions that varied only with respect to t_{\max} , the ordering of responses was in no systematic way related to the values of t_{\max} . Since inspecting a longer period of the same process should lead to a rightward shift on the response scale, such systematic inconsistency can only mean that the S does not understand the task, or that the criterion is shifted all the time. Omission of these Ss does not artifactually induce an effect of t_{\max} , as it is quite possible that the remaining Ss responded consistently, but with the same underestimation for all values of t_{\max} . The results of the 23 consistent Ss are presented in fig. 2b. For these Ss the effect of t_{\max} is highly significant ($F_{2, 44} = 8.44, p < 0.001$) but the interaction with b ($F_{4, 88} = 3.26, p < 0.05$) means that the advantage of inspecting a larger proportion of a process is rapidly lost when b increases. The effect of d (fig. 3a) was not significant ($F_{3, 105} = 2.35, p > 0.05$) nor was the interaction with b ($F_{6, 210} = 1.63, p > 0.10$). As the interaction with Ss was highly significant ($F_{105, 420} = 2.46, p < 0.001$) the results of the 23 consistent Ss were again analyzed separately. For these Ss (fig. 4b) both effects were significant ($F_{3, 66} = 3.32, p < 0.05$; $F_{6, 132} = 3.31; p < 0.01$). However, a *post-hoc* Newman-Keuls analysis (Winer 1962) revealed that the only effect of duration occurred between the duration of 60 sec and the others, and only for $b = 0.1$. The interaction between t_{\max} and d did not reach significance ($F_{6, 210} < 1$) not even for the consistent Ss ($F_{6, 132} < 1$).

Discussion

The effect of underestimation of exponential growth is clearly demonstrated in conditions where presentation is non-numerical and spread over time. The size of the effect depends on the value of b ; the effect of t_{\max} (the proportion inspected before the extrapolation) is only significant when inconsistent Ss are excluded.

The marked overestimation of growth that occurred in the condition $b = 0.1$ might partially reflect the Ss ' need to use the total response scale; since the majority of the marks fell in the left half of the scale the Ss might have been seduced to use the right-hand half in those conditions that were the least prone to underestimation.

What kind of model could account for these data or, in other words, what kind of strategy could Ss employ? When they take into account only the most recent situation as it is represented by n and t_{\max} , and extrapolate from thereon in a linear manner, the values of β would be very close to zero, independently of b , t_{\max} and d . Since this is not the case it is obvious that the Ss noticed that the growth was not linear. This is further illustrated when the conditions ($b = 0.3, t_{\max} = 0.50 d$) and ($b = 0.5, t_{\max} = 0.67 d$) are compared. In both conditions the display is stopped at $n = 47$ (18.4%). On the average the marks on the response scale were placed at respectively 41% and 50%, thus reflecting the impact of the non-linear components.

Two distinctly differing strategies could lead to some sort of growth estimation; we will call these strategies *time sampling* and *quantity sampling*. In time sampling the process is sampled after $x, 2x, 3x \dots$ etc. updatings; the basic information then consists of a series of ever shortening time intervals. In quantity sampling the process is sampled at regular time intervals; the basic information then consists of a series of ever increasing quantities.

In the two following experiments it is attempted to discriminate between these strategies by manipulating the frequency at which the display is updated. In experiment 2 (fig. 4a) the display is updated every 1, 2 or 4 duckweeds. We will call this variable *blocksize*. The information contained in every updating is the time interval between the updatings. When Ss employ a time sampling strategy the effect of increased blocksize is a reduction of information along the whole trajectory. When, instead, Ss employ a quantity sampling strategy the effect of increased blocksize is quite detrimental at the beginning of the presentation, but small at the end; this variable effect of blocksize will depend very much on the values of b and t_{\max} . Hence, if manipulation of blocksize interferes with the Ss' strategy at all, an interaction with b and t_{\max} would indicate a quantity sampling strategy.

In experiment 3 (see fig. 4b) the display is updated at regular time intervals. Now the effects of the two sampling strategies become reversed: if Ss use a time sampling strategy they will get more and more deprived of information as the process goes on, dependent on the values of b and t_{\max} . Hence, if reduction of the number of updatings has any effect at all, an interaction with b and t_{\max} would indicate a time sampling strategy.

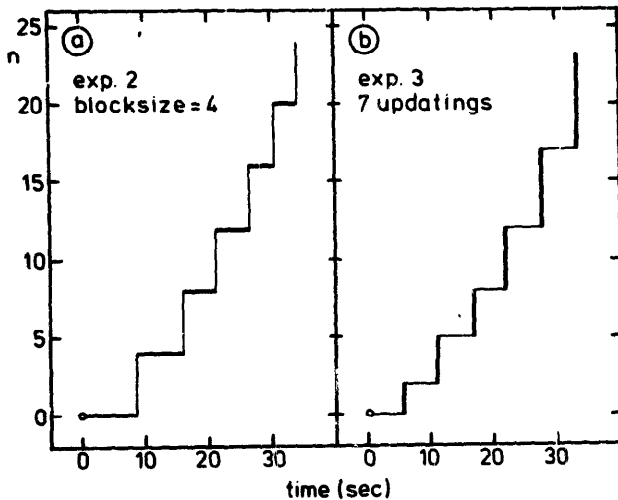


Fig. 4. Updating in the condition $b = 0.3$, $d = 100$ sec, $t_{\max} = 0.33d$, as limited by the procedures of experiment 2 and 3.

Experiment 2

The basic set-up and the choice of parameters were the same as described before. Duration (J) was excluded as an independent variable; instead it was randomly varied between 100 and 200 sec, in order to exclude the possibility that the correct response is obtained by responding in proportion to presentation time. The new element in this experiment was that updating occurred every 1, 2 or 4 duckweeds. The number of updates (including the empty pond as a starting configuration) varied from 3 ($b = 0.5$, $t_{\max} = 0.33d$, blocksize = 4) to 145 ($b = 0.1$, $t_{\max} = 0.67d$, blocksize = 1). Increasing blocksize still further was not possible as the number of updates should never be below 3.

Method

Apparatus and stimuli

Apart from the variable blocksize, the stimulus situation was as in experiment 1.

Subjects

The S s were 31 male and female students from the State University of Utrecht. They were paid Dfl. 25,- for the completion of the experiment.

Procedure

Factorial combination of the variables b , t_{\max} and blocksize yields 27 conditions. All S s worked through the 27 conditions twice in three sessions of 18 trials. Blocksize was only changed between sessions. The order of presentation within sessions was random; assignment of S s to session orders was also random. In order to reduce the frequency of inconsistent responses, the S s' understanding of the instruction was checked carefully.

Results

An overview of the results is presented in fig. 5. First it should be noticed that the results resemble those presented in fig. 2b. The effects of b , t_{\max} and their interaction were highly significant ($F_{2, 60} = 25.70$, $p < 0.001$; $F_{2, 60} = 34.90$, $p < 0.001$; $F_{4, 120} = 4.83$, $p < 0.01$). No S s showed marked and systematic inconsistencies. The main effect of blocksize was insignificant ($F_{2, 60} < 1$). First and second presentation of the processes was not a source of variation ($F_{1, 30} < 1$). The presentation of all blocksize conditions to all S s might have introduced a range effect (Poulton 1975) thus obscuring a possible effect of blocksize. This was not the case, however, as is shown from a separate analysis of first sessions only: the effect of blocksize was still insignificant ($F_{2, 21} < 1$). As explained before, blocksize, though not significant as a main effect, might interact with b and t_{\max} when S s use a quantity sampling strategy. Neither the first order interactions between b and blocksize, t_{\max} and blocksize, nor the second order interaction between b , t_{\max} and blocksize, were significant however ($F_{4, 120} < 1$; $F_{4, 120} < 1$; $F_{8, 240} < 1$).

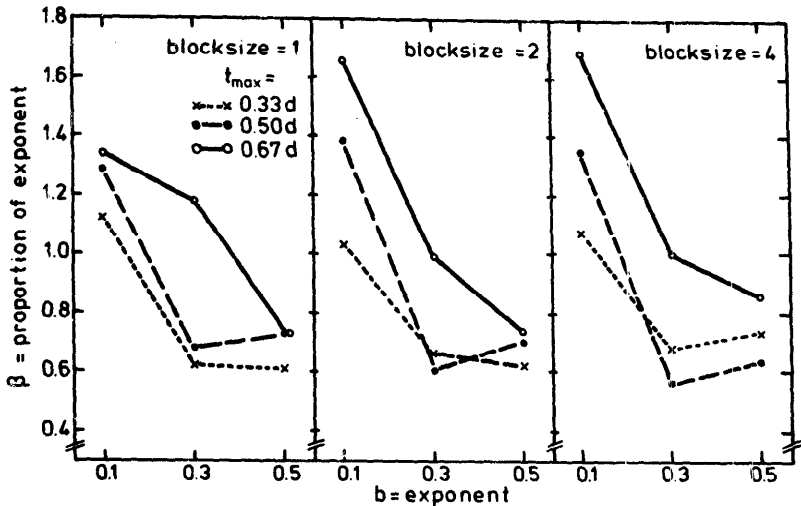


Fig. 5. Results of experiment 2; effects of blocksize and t_{\max} .

Discussion

The main conclusions of experiment 1 are confirmed. Concealment of the microstructure of the process by increasing blocksize did not have any effect on the β -scores; obviously *Ss* do not use that sort of detailed information. Since no effects of blocksize were observed we cannot decide between time sampling and quantity sampling strategies.

The information used by the *Ss* must be very rough indeed, as even in the conditions $b = 0.5$, $t_{\max} = 0.33d$ no effect is observed when the number of updatings is reduced to three only (see fig. 7a). The conclusion can be drawn that *Ss* either use a time sampling strategy or a quantity sampling strategy based only on the beginning, the middle and the end of the displayed growth functions.

Experiment 3

The idea of this experiment was to reduce the information presented to the *Ss* by presenting 3, 5 or 7 data points with equal time intervals. The presentation of the empty pond was counted as the first updating. Example: in the case of three updatings and $t_{\max} = 0.50d$ updating occurred at $t = 0$ sec, $0.25d$ and $0.50d$. As in experiment 2, d was randomly chosen from the open interval 100 to 200 sec; the time interval between updatings varied from 5.6 sec (7 updatings, $t_{\max} = 0.33d$, $d = 100$ sec) to 66.7 sec (3 updatings, $t_{\max} = 0.67d$, $d = 200$ sec).

Method

Apparatus and stimuli

Apart from the variable number of updatings, the stimulus situation was as in experiment 1.

Subjects

The Ss were 30 male and female students from the State University of Utrecht. They were paid Dfl. 25,- for the completion of the experiment.

Procedure

Factorial combination of the variables b , t_{\max} and number of updatings yields 27 conditions. All Ss worked through all 27 conditions twice in three sessions of 18 trials. Number of updatings was changed only between sessions. The order of presentation within sessions was random, and so was assignment of Ss to session orders.

Results

The results are presented in fig. 6. Again the effects of b , t_{\max} and their interaction were significant ($F_{2, 58} = 40.6, p < 0.001$; $F_{2, 58} = 3.91, p < 0.05$; $F_{4, 116} = 5.47, p < 0.001$). The $b \times t_{\max}$ interaction shows again that the advantage of long inspection times disappears when the growth gets faster.

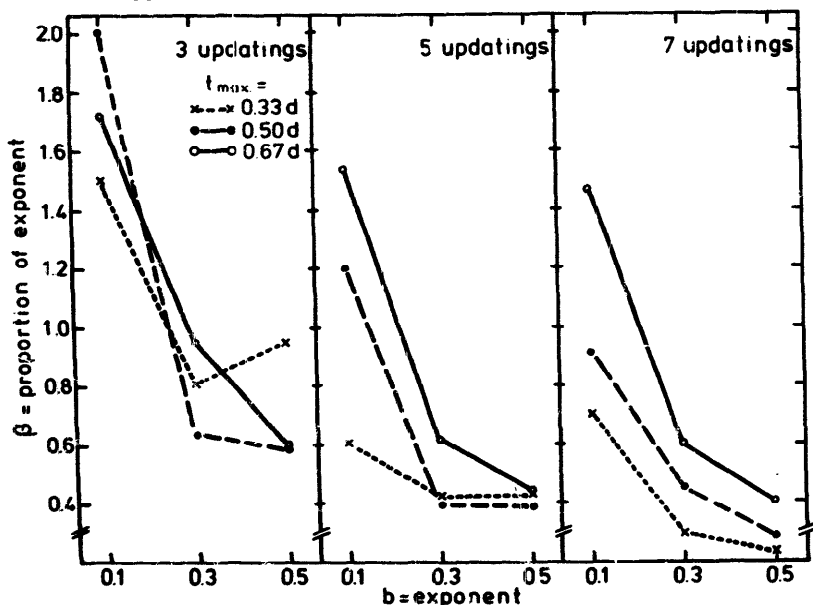


Fig. 6. Results of experiment 3; effects of number of updatings and t_{\max} .

The main effect of number of updatings was highly significant ($F_{2, 58} = 9.08$, $p < 0.001$). The average scores were $\beta = 1.09, 0.67, 0.59$ for 3, 5 and 7 updatings respectively, against 0.86 in experiment 1 and 0.92 in experiment 2. The interactions of this effect with b , t_{\max} and $b \times t_{\max}$ never reached significance ($F_{4, 116} = 1.55$, $p > 0.10$; $F_{4, 116} = 1.09$, $p > 0.30$; $F_{8, 232} = 1.75$, $p > 0.05$). First vs. second presentation of the stimuli again did not yield a significant difference ($F_{1, 29} < 1$).

Discussion

The results show that underestimation of growth becomes worse when the number of updatings increases from 3 to 7: the more information, the worse the extrapolation. This effect can also be observed in every-day life when changes are more readily detected by people that have been away a long time. The absence of significant interactions with b and t_{\max} suggests that the Ss used a quantity sampling strategy.

How could it happen that the scores of experiment 3 were lower than in the previous experiments when the number of updatings increased? One possible explanation is that in experiment 1 and 2 Ss always selected the same sampling rate, irrespective of the variable blocksize. In experiment 3, however, samples are presented at regular intervals and that is the way Ss tend to sample. Consequently they might adopt the sampling rate imposed by the experimenter.

From experiment 2 we knew already that the sampling rate voluntarily chosen by the Ss might be low. On the assumption that the effect of updating rate is to influence the sampling rate chosen by the Ss we may interpret fig. 7a to mean that the natural sampling rate chosen by an S is about 3 or 4. Is this number independent of t_{\max} ? If not, one would predict the number of samples used in the condition $b = 0.5$, $t_{\max} = 0.67d$ to be 6 or 8. Fig. 7b suggests that the sampling rate in the latter condition is still below 6, and presumably 3 or 4. Thus it is indicated that the number of samples taken is unrelated to the proportion of the process that is

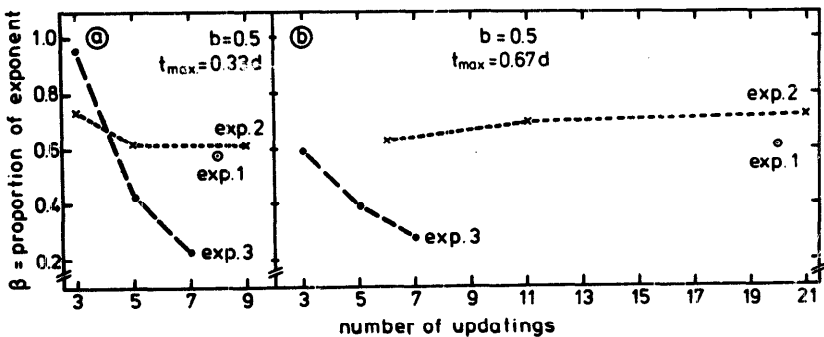


Fig. 7. A comparison across three experiments. Reduction of the number of updatings had a marked effect in experiment 3, and not in experiment 2.

presented. As the *S* does not know this proportion in advance, samples must be chosen from memory, after presentation of the growth process. What will actually happen is that the *S* monitors the process until the presentation stops, then three quantity samples are taken (the beginning, which is just the empty pond; one sample somewhere from the middle; the end which is the number of duckweeds last shown) and then the function describing these three data points is extrapolated.

The descriptive value of this model may appear from the following thought experiment. Take three samples from the condition $b = 0.5$, $t_{\max} = 0.50d$: 0 (the beginning), 4 (n at $t = 0.25d$) and 19 (n at $t = 0.50d$). If we extrapolate by keeping the difference of differences constant (that is: fitting the quadratic polynomial) we reach $n = 256$ after 3.5 times the display time; the subjective estimate would be that we are at $1/3.5$ is 29% of the total time needed to get the pond filled. The point $t = 0.29d$, $n = 19$ lies on an exponential function (Equation 1) with $b = 0.3$. The exponent is therefore weighed with $\beta = 0.3/0.5 = 0.6$ which coincides closely with the experimental data for the condition $b = 0.5$. In a similar way we obtain $\beta = 0.9$ for $b = 0.1$ and $\beta = 0.06$ for $b = 1.0$ (which fits nicely with the data presented by Wagenaar and Sagaria).

As a logical consequence of imposed higher sampling rates *Ss* are confronted with smaller differences between successive samples. When *Ss* judge growth on the basis of differences it is quite possible that the small differences occurring when data points are interpolated obscure the growth. This could be brought about by some sort of threshold mechanism: if higher-order effects (differences of differences, and so on) are too small they might be omitted. Higher-order effects are always smaller when the sampling becomes more fine-grained.

Some supporting evidence was presented by Timmers and Wagenaar (1977) who showed that decreasing exponential functions are much better extrapolated than increasing ones. This difference provides some additional support for our conclusion that *Ss* employ a quantity sampling strategy. The result of quantity sampling and time sampling are respectively a series of ever increasing quantities and a series of ever decreasing time intervals. It would be hard to understand the large effects of underestimation observed in the present experiments, if *Ss* extrapolate on the basis of a decreasing series of time intervals.

General conclusion

The underestimation of exponential growth reported before is not limited to situations in which the process is presented by means of tables and graphs; when the process is presented as it develops in time, subjects underestimate the growth to the same extent, except when the growth is almost linear.

It helps a little when a larger part of the process is inspected by the subjects, but the advantage disappears quickly when the exponent increases above 0.1.

The experiments on variable blocksize and updating frequency suggest that subjects use a quantity sampling strategy and that they use only three samples for extrapolation of the process.

References

- Poulton, E. C., 1975. Range effects in experiments on people. *American Journal of Psychology* 18, 3–32.
- Price, D. J. de Solla, 1963. *Little science, big science*. New York: Columbia University Press.
- Timmers, H. and W. A. Wagenaar, 1977. Inverse statistics and the misperception of exponential growth. *Perception and Psychophysics* 21, 558–562.
- Wagenaar, W. A., 1969. Note on the construction of digram-balanced latin squares. *Psychological Bulletin* 72, 384–386.
- Wagenaar, W. A. and S. D. Sagaria, 1975. Misperception of exponential growth. *Perception and Psychophysics* 18, 422–426.
- Wagenaar, W. A. and H. Timmers, 1978. Intuitive prediction of growth. In: D. F. Burkhardt and W. H. Ittelson (eds.), *Environmental assessment of socioeconomic systems*. London: Plenum Publishing Corporation.
- Winer, B. J., 1962. *Statistical principles in experimental design*. New York: McGraw-Hill.