N86-32998

HItts' Law? A Test of the Relationship Between Information Load and Movement Precision

Mathew Zaleski
SWS Systems

Neville Moray
Department of Industrial Engineering University of Toronto

Recent Technological developements have made viable a man-machine interface heavily dependent on graphics and pointing devices. This has led to new interest in classical reaction and movement time work by Human Factors specialists.

Two experiments were designed and run to test the dependence of target capture time on information load (Hitt's Law) and movement precision (Fitts' Law). The proposed model linearly combines Hick's and Fitts' results into a combination law which then might be called Hitts' Law. Subjects were required to react to stimuli by manipulating a joystick so as to cause a cursor to capture a target on a CRT screen. Response entropy and the relative precision of the capture movement were crossed in a factorial design and data obtained that were found to support the model.

Introduction

Software engineers have always been under pressure from software users to provide friendly, easy to use interfaces to software systems of all kinds. The most effective seem to be those which combine with hardware to enable the user to point out places on the computer screen. Several systems are on the market which use a mouse for this purpose.

Briefly, the software systems of which we wish to make an example of are those that have come to be called "icon driven". A typical example is the Finder of Apple Macintosh. It is relevant to note that the users of these systems make commands to the computer by pointing to small pictures on the screen. The user's progress is sometimes limited by the speed with which these "icons" can be selected .

Human factors engineers have undertaken to study the properties of several of the pointing devices. Card, English and Burr [1978] demonstrated that the mouse and joystick are limited by the classical psychological result of Fitts [1954]. Further work made clear that using a joystick to control the motion of a cursor on a CRT is subject to the same fundamental limitations as are manual aimed motions, say with a stylus. Studies in this field typically ask the subject to manipulate a mouse or joystick so that a computer controlled cursor moves within a target area of the computer's CRT. The time required for the subject to make movements of varying length towards targets of varying size is measured (MT), and usually found to follow the well known result:

$$MT = a\,ID + b, \qquad (1)$$

where ID is called the Index of Difficulty and is usually:

$$ID = \log_2 (2A/W). \qquad (2)$$

Relation 1 is called Fitts' Law, and equation 2 is only one definition of ID. Many others have been proposed, for instance in Welford [1968].

Fitts Law is closely related to information theory. There is no derivation of FItts Law in the rigorous sense, but fairly convincing analogies can be made which compare movements made by the human to transmitting information down a noisy channel. Consider a user about to make a cursor motion. It is intuitive that he is able to transmit more information with a precise movement than a crude one. If we further suppose man's motor system has a finite capacity to transmit information, then we expect that the time required to execute a motion ought to be proportional to the amount of information transmitted. Given that ID measures the information content of a motion, equation 1 follows.

There is another important element of the user's task, namely that he must often choose between discrete alternatives that are clearly presented on the screen before him. In many cases he is performing a similar task to that performed by the subject of a choice reaction time experiment. (Hick [1952], Hyman [1953])

The information content of a discrete target capture is quantified by a measure called response entropy (H). If we assume that man has a limited capacity to transmit information then we conclude that the length of time to make a choice (RT) will depend on the entropy of the required response.

$$RT = c + d\,H \tag{4}$$

Equation 4 (Hick-Hyman Law) is often written for equiprobable stimuli (so the probability of each is 1/n, where n is the number of stimuli) as:

$$RT = c + d\,log_2(n). \tag{5}$$

In light of the above discussion one might remark that there are aspects of the icon driven software interface which correspond to the view of both the Hick-Hyman and Fitts' Laws. A natural question to ask is whether a combination of the two laws might not be a useful way of modelling the behaviour of the user of such software. One might suggest that time taken to capture one of several targets would be described by:

$$CT = \alpha + \beta\,H + \gamma\,ID \tag{6}$$

where CT is the time to capture the target, H is the average response entropy and ID is the index of difficulty of the movement.

This combination of Hick's and Fitts' laws was proposed by Beggs, Graham, Monk, Shaw and Howarth [1972]. They performed an experiment which had inconclusive results and so it would appear that such a combination law has never been proved. If the combination law were found to hold it would offer a more complete model of the operator of icon or menu driven software systems in that it would incorporate two aspects of performance, namely the effects of both movement precision and response entropy on the average capture time.

In suggesting a additive combination of two fundamentally different psychological processes one enters a Great Debate in modern psychology. If a combination law such as equation (6) is found to hold does this imply that the underlying internal processes are serial and additive? Sternberg [1969] performed an elegant series of experiments in which certain memory searching processes appeared to be carried out in a highly serial way. His work gave rise to what has come to be called the additive factors methodology, which once was viewed as a way of detecting serial vs parallel processing. Taylor [1976] amongst others, suggested parallel processing schemes to explain the same data, and hence introduced a more conservative experimental approach which, unfortunately, is much more complex. This is mentioned in the context of this study because the data in the present study were analysed in a way similar to that used in additive factors and thus the results may be interpreted accordingly.

Goals of the experiments

The immediate goal of the study was to test whether the combination law holds for the task of manipulating a joystick so that in response to a visual stimulus a computer

SCREEN LAYOUT

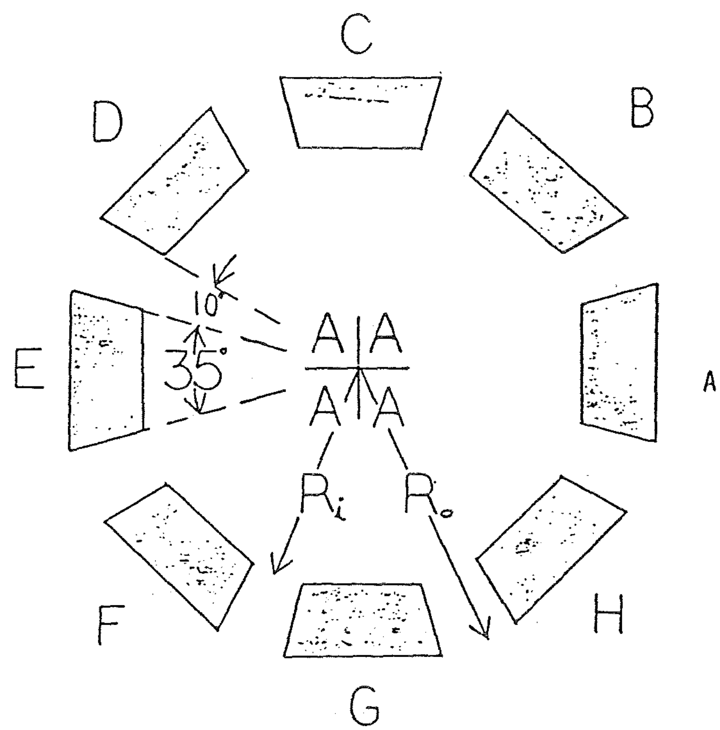


fig 1

controlled cursor moves into a target area of a CRT display. A positive result would open the door to a predictive model which might be useful as part of the design process.

Description of pilot experiment

The pilot experiment set out to test equation 6 in a very typical choice reaction time study. The experiment was carried out on a micro computer with subjects manipulating a joystick and the computer arranging that a cursor move correspondingly. Reactions were to letters of the alphabet plotted in the centre of the screen. There was a one to one correspondence between targets and letters. In a given trial the number of targets and their dimension was manipulated. Stimuli were chosen such that they were equiprobable, so the average response entropy of a sequence of trials with n targets was $log_2(n)$. All the targets in a given trial were of the same relative size, and given the geometry of the situation the index of difficulty or ID was given by:

$$ID = log_2[\ (R_o + R_i)/(R_o - R_i)\ ] \qquad (6)$$

where R_o and R_i are defined in figure 1.

The pace of the experiment was sedate. First the targets and cross-hairs for a particular trial were plotted on the CRT. The subject was verbally instructed to centre the cursor on the cross-hairs during a delay of about 3.4 sec. An auditory warning followed the delay and 4 copies of the stimulus were presented (as in figure 1). Subjects then captured the target with the instructions to be as time efficient as possible without sacrificing accuracy. The cursor remained under joystick control for a further 3 seconds following the onset of the stimulus, at which time the screen cleared and there was a one second delay while data was written to disk. There was no quantitative feedback at the end of the trial or session of the subjects' performance. No attempt was made to instill competition between the subjects.

Experimental Design

There were three levels of H (see table 1) and three levels of ID (table 2). Subjects performed four pairs of sessions, each pair constituting one pass through the design. In most cases both halves of the design were performed one shortly after the other, with a rest in between. Each session was composed of six "blocks" of trials. All trials within each block had the same number of targets and hence the same response entropy. Each block was in turn divided into three "groups" of six trials each. All trials within a group had constant ID. Thus each session was made up of 12 trials in each of the nine cells of the design. Since it took two sessions for one pass through the entire design, each pass required 24 trials in each cell for 216 trials in all. Seven subjects, who were graduate students at the University of Toronto, each completed four passes through the experiment . Subjects were paid $5 per hour for their participation in the experiment.

Table 1 - H manipulation - pilot

response entropy (H)	number of targets	targets sets	special features
3	8	ABCDEFGH	all targets used
2	4	ACEG BDFH	vert/Horiz quadrants diagonal quadrants
1	2	AE BF CG DH	right and left diagonal 180 ° apart top and bottom like BF, but rotated 90 °

Table 2 - ID manipulation - pilot

ID	R_t	R_o
3	60	77
4	75	85
5	80	85

In the blocks which contained fewer than all the targets there was the problem of choosing which subset of targets to use. The subsets chosen are listed in table 1. Note that what has been done is to restrict the screens to either vertical and horizontal or diagonal symmetry but no mixture of the two.

Implementation

The experiment was run on an Apple IIe 6502 based micro computer. All software was written in UCSD pascal except the clock and ADC drivers, which were written in assembler. A real time clock and ADC device handler was designed which collected data while the pascal mainline controlled the screen.

Subjects responded using a Measurement Systems joystick with no spring return to centre. The maximum possible deflection of the joystick was about 30°. Subjects were not located exactly with respect to the screen and joystick, but for the typical subject there was a gain of about 0.25° of visual angle for each 1° of joystick deflection.
With this apparatus the duration of each target capture (CT) was defined to be the interval between the onset of the stimulus and the beginning of a 350 millisecond capture of the target. Reaction Time (RT) was defined as the period from the onset of the stimulus until the joystick was deflected 0.3°. Movement time (MT) was the difference between CT and RT.

Results of Pilot Experiment

Statistical analysis was carried out in two main ways using analysis of variance (ANOVA) and regression analysis. This reflects the three main topics of interest, which are:

i. How well do the independent factors ID and H predict the time required by subjects to select and execute a target capture response?

ii. In terms of the information hypothesis, what are the information capacities of the subjects to H and ID? Do they change with practice?
iii. Do the independent factors interact?

Results of ANOVA analysis for pilot experiment

An ANOVA was carried out for RT, MT and CT for each of the four runs the subject made through the experimental design. The ANOVAs assumed a three factor completely randomized mixed model. Only the results of the last run will be described here. Note however that examination of the training data showed that subjects were not stable at the end of the pilot experiment, as they were still improving significantly from the third to the fourth and last session. In the following ANOVA data the standard statistics are presented, namely the F score of the null hypothesis test (F), its Mean Square Error (MSE), the probability of the null hypothesis being true (p) and finally the fraction of the variance attributable to each factor, ω^2.

Reaction Time

The ANOVA shows that RT is influenced significantly only by H. Interaction terms and ID have no significant effect. Between subjects variation accounts for much of the total variance. The difference between subjects is highly significant.

We emphasize that RT has been defined operationally to be the time from the onset of the stimulus until the first small deflection of the joystick. Thus all factors which cause the subject to delay are grouped under RT. Nevertheless, as shown by the ANOVAs described below, only H has a significant effect on RT. This would imply that the particulars of the movement about to be made do not affect the duration of the delay before the movement.

Table 3 - Reaction Time Anova - pilot

factor	F		MSE	p	ω^2
Subject	F(1,6)=	79	2830	< 0.001	-
ID	F(2,12)=	1.2	22	0.30	-
H	F(2,12)=	32	99.6	< 0.001	0.24
interaction	F(4,24)=	.98	13	0.44	-

Movement Time

Movement time was found to depend significantly only on ID. H and interaction terms were found to be non-significant. As before subjects differed significantly.

Table 4 - Movement Time Anova - pilot

factor	F		MSE	p	ω^2
Subject	F(1,6)=	251	760	<0.001	-
ID	F(2,12)=	82	73	<0.001	0.60
H	F(2,12)=	0.062	31	0.94	-
interact	F(4,24)=	.53	66	0.71	-

Capture Time

Capture time showed significant effects of H and ID, but the interaction component of the model was not significant. The ω^2 column shows how ID accounts for about 20% of the variance and H for slightly more than 10%. The rest is due to between subject differences. It is clear from this data that there is no interaction taking place between the factor which affects the period of time until the beginning of the subjects' overt response (H) and the factor which affects the duration of the movement (ID).

Table 5 - Capture Time Anova - Pilot

factor	F		MSE	p	ω^2
Subject	F(1,6)=	173	4542	<0.001	-
ID	F(2,12)=	63	76	<0.001	0.20
H	F(2,12)=	20	166	<0.001	0.12
interact	F(4,24)=	3	73	0.53	-

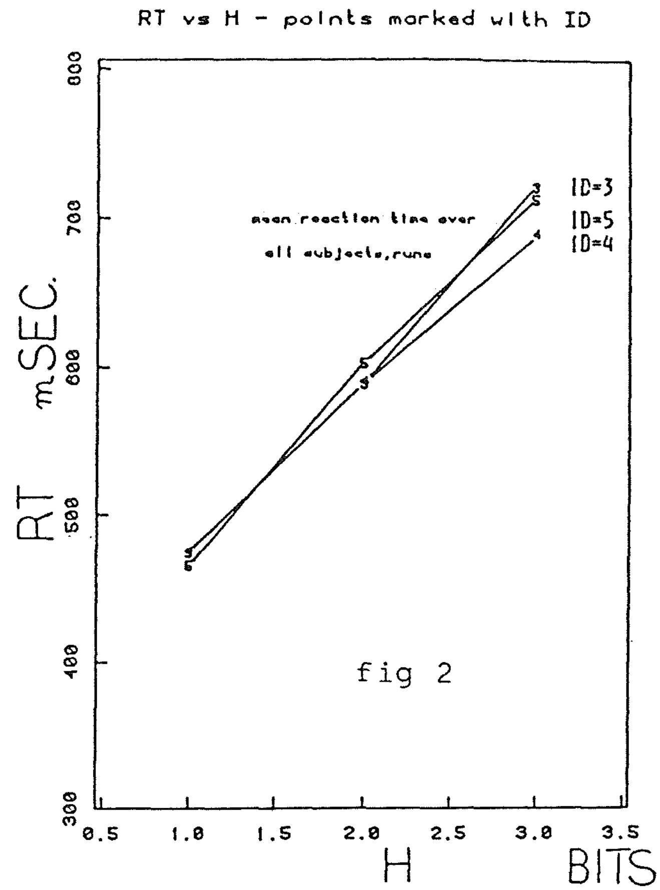
Results of multiple regression analysis - pilot study

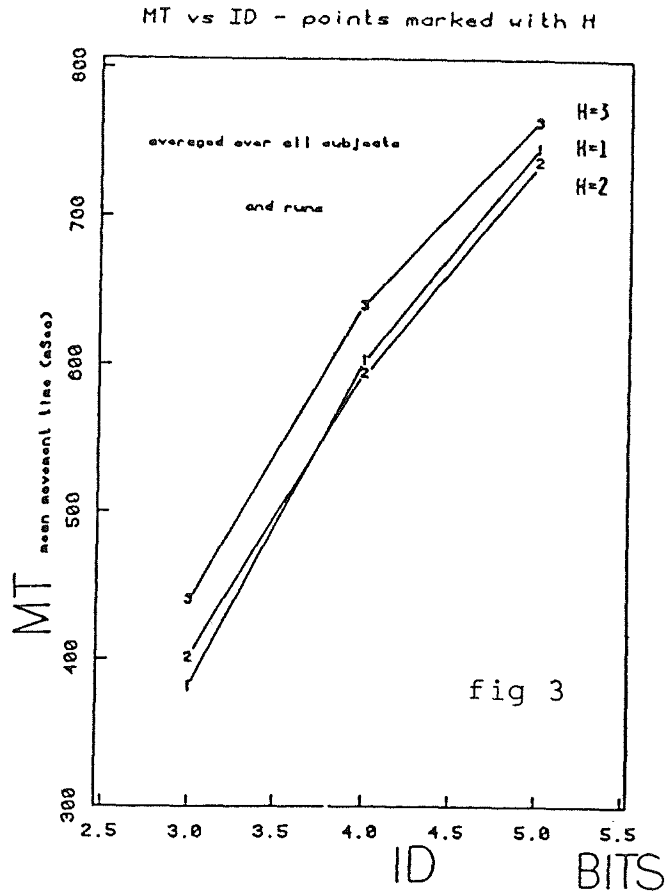
The result to be presented in this section is the multiple regression of average (across all subjects) CT vs ID and H. This regression is an experimental test of the combination of Fitts' Law and the Hick-Hyman Law given in equation 6. Simple regressions of RT vs H and MT vs ID were also performed to continue our examination of how these stages depend on H and ID. In this section we will concentrate on two main statistics:

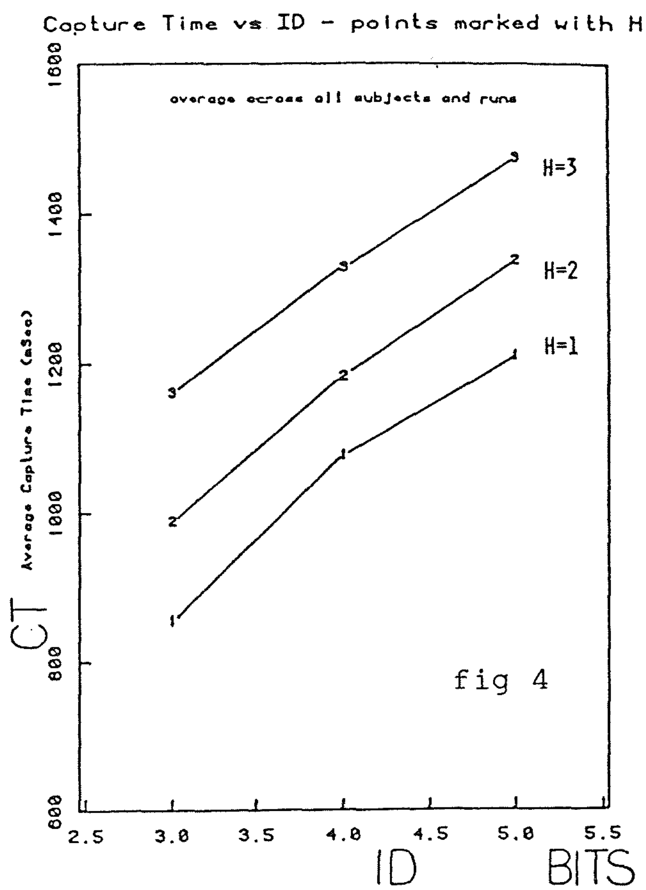
i. The statistic r^2 is quoted to describe what proportion of the total variance can be explained using the simple linear model of equation 6. It is generally not safe to judge the quality of fit from r^2 alone.

ii. Residual Standard Error (RSE) is equal to the average square residual. This gives an indication of how far the average data point is from the fitted line.

The coefficients of the regression have the physical dimension of seconds per bit. Thus their reciprocal has dimensions of information capacity, or bits per second. The coefficients can be used to calculate the information capacity of the subjects with respect to H and to ID. Comparisons of the resulting information capacities to those of earlier studies is discussed in later sections.


RT vs H - points marked with ID
mean reaction time over
all subjects,runs
ID=3
ID=5
ID=4
RT mSEC.
H
BITS
300
400
500
600
700
800
0.5
1.0
1.5
2.0
2.5
3.0
3.5
fig 2


MT vs ID - points marked with H
averaged over all subjects
and runs
H=3
H=1
H=2
MT mean movement time (mSec)
ID
BITS
300
400
500
600
700
800
2.5
3.0
3.5
4.0
4.5
5.0
5.5
fig 3


Capture Time vs ID - points marked with H
average across all subjects and runs
H=3
H=2
H=1
CT Average Capture Time (mSec)
ID
BITS
600
800
1000
1200
1400
1600
2.5
3.0
3.5
4.0
4.5
5.0
5.5
fig 4

Reaction time

This section presents the extent to which the data follows the Hick-Hyman Law.See figure 2 for a graph of RT vs H for data pooled across all subjects' last run.

Table 6 - Reaction Time regression analysis - pilot

for model: $RT=\alpha+\beta H$
(Hick's Law)

α	β	r^2	RSE	F(1,7)
349	123	0.98	14.7	421
mSec	mSec/bit			

Movement time

The Fitts' Law component of the data is described in Table 7. Presumably the value of the constant term would be different if the arbitrary boundary between RT and MT were changed.

Table 7 - Movement Time regression analysis - pilot

for model: $MT=\alpha+\gamma ID$
(Fitts' Law)

α	γ	r^2	RSE	F(1,7)
-122	168	0.98	19.9	428
mSec	mSec/bit			

Capture time

The test of equation 6 with respect to the data of the pilot experiment is presented here. The r^2 of the regression is 0.99, so the model is explaining almost all of the variance in the pooled data. It is fair to say that the combination law describes the CT data just as well as the two classical laws describe RT and MT. Table 8 summarizes the regression results for the average across all subjects fourth run. Each subjects has made 648 responses previously.

Table 8 - Capture Time regression alalysis - pilot

α	β	γ	r^2	RSE
264	127	150	0.99	21
mSec	mSec/bit			

One might ask how such high linearity is present given the large between subjects variation measured by the ANOVAs. Table 9 presents the same regression carried out in Table 8 except for each subject individually. Since there is less data there is more noise. Initially it was anticipated that the subjects would differ mostly in intercept with roughly similar coefficients. This, as is shown in table 9, is not borne out by the data at all. The per subject regressions illustrate again how simple models can predict mean behaviour well and yet cast little light on individual performance.

Contest - study 2

In the task of the pilot study the additive combination of Hick's and Fitts' laws was a very appropriate way of mathematically describing average subject performance. The results indicated that the subjects could be thought of a reacting in two sequential

Table 9
Per Subject Capture Time regression alalysis - pilot

model: $CT=\alpha+\beta H+\gamma ID$				
subject	α	β	γ	r^2
gc	203	107	152	0.82
ir	279	52	154	0.84
jp	149	103	219	0.87
kh	253	170	130	0.85
kv	321	122	112	0.88
mc	94	78	160	0.81
mk	549	257	127	0.94
average	264	127	150	.99
dimension	mSec	mSec/bit	mSec/bit	-

phases: a response selection stage followed by a movement stage. The task carried out by subjects in the pilot experiment differed from otherwise similar tasks performed by operators of icon driven software systems in several important respects:

i. The trials were highly discrete. There was a gap between trials which was of considerably longer duration then the trials themselves. The experiment of Beggs et al [1972] was a continuous one and H and ID were found to interact. Practical software systems often require the user to make a series of captures, and often with little or no externally imposed temporal uncertainty.

ii. The symmetry of the target capture motions made in the pilot experiment was highly radial. The direction of the required motion corresponded one to one with the stimuli. This is artificial in the sense that in practical situations the stimulus corresponds to a target, but the direction of motion depends upon the starting position as well.

iv. No feedback was given to the subjects of the pilot study of when they had captured their target (other than the position of the cursor on the screen) or how their performance compared to other subjects. This is very unrealistic, for in a practical setting there is little point in capturing targets if nothing is going to happen when you do so.

SCREEN LAYOUT

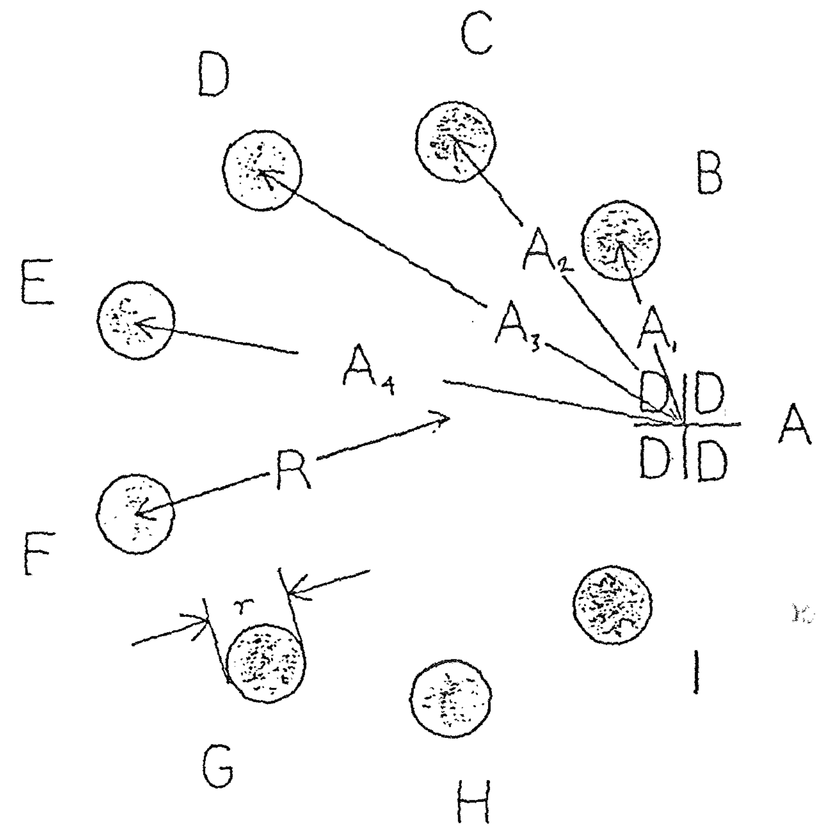


fig 5

Description of Contest

The second study was designed in answer to these points. Enough feedback was built into the experiment that we refer to it as "Contest". Trial by trial feedback was provided by sounding a beep as soon as the capture was detected. At the end of each session a subject was told his total score, and a graph of these scores was clearly displayed on which each subjects progress was recorded in a different colour. Much effort was expended preparing software to make it possible to make the end of each session a competetive event, in which scores were compared (and excuses made). A prize of $10 was offered (above the hourly rate) for the best score.

Contest was continuous, which in the context of such experiments means only that delays between the trials have been minimized. The subject no longer waits for the stimulus, but initializes its onset himself by completing the previous trial. Subjects face no temporal uncertainty apart from that produced by short delays in the software. Hopefully this will allow more direct comparison to the continuous tapping style of Fitts Law experiment [Fitts 1954]. The response motions no longer correspond only to stimuli but also depend on the situation. In Contest the subjects do not return to cross-hairs at the centre of the screen but rather the cross-hairs are plotted over the last target captured, so that the next trial can start immediately. Figure 5 illustrates the screen layout of Contest.

It is desirable to use an Index of Difficulty as comparable to the one used in the pilot study as possible. Referring to figure 5 we note:

$$\text{ID} = log_2\,(A_i/r). \qquad (7)$$

Experimental Design - Contest

As is visible from figure 5, up to four IDs are introduced by one choice of R and r. This has the effect of making it impossible to separate trials into groups of constant ID. For some "configurations" (figure 5 is an example of one configuration) there can be trials of different ID. The experimental design was difficult because it was convenient to have the same number of trials in each cell. However, we could not choose simple subsets of the targets as in the pilot study (see table 3.1.1) because there was no such set which had the same number of trials in each cell. The result was that a relatively large number of different configurations were chosen.

There were six subjects, three men and one woman graduate students, and two high-school age teenagers. They were paid $5 per hour and knew about the $10 first prize from the outset. Subjects participated in 11 or 12 sessions, until their behaviour had asymptoted as indicated by their score in each run.

There were 12 levels of ID and 4 levels of H; 48 cells in all. The design was made up of 32 blocks of three groups of 16 trials each for 3x16x32 = 1536 trials. Runs 1 through 10 used exactly the same sequence of trials, then for the last two trials the sequence was changed. The block and group structure was identical, but the stimuli were presented in a different order.

Results of Contest

The implementation of the experiment was essentially the same as for the pilot study, except for some rearrangement of the screen. The data processing had to be streamlined in order to detect target hits on line and to be ready at the end of a trial to feed performance back to the subject.

The statistical processing applied was also unchanged. The only difference was that it was found that the division of CT into RT and MT was not possible and so the statistics are quoted only for CT. In the pilot study there was a long forewarning period in which the subjects kept the cursor still on the cross-hairs. Thus the end of the Reaction Time period was reliably detected by the first deflection of the joystick. In Contest, continuous by design, subjects never held the joystick still for long enough to detect any transition between RT and MT. One attempt was to estimate the rate and acceleration of the cursor and make a decision based on them, but the results were not encouraging.

A session consisted of 1536 captures, the average trial requiring about 0.6 sec each, for an average session duration of about 50 minutes. Since the pace was set mostly by the subjects the percentage of the time on task actually spent in control of the cursor was about 60%. Subjects found the sessions quite tiring. In retrospect, a session of about 1000 trials would have been more appropriate.

Results of ANOVA analysis - Contest

ANOVA showed that variance in CT data pooled across all subjects' most highly trained session was almost entirely explained by the factors H and ID. H was responsible for 44% of the total variance, ID for 48%, leaving very little for between between subject differences and interaction. The interaction of H and ID was not significant at the 2.5% level, but it was at the 5% level. However, if the ω^2 of the interaction term is examined it becomes clear that the interaction has a negligible effect on CT. See table 10. It is fair to say that the interaction, even though statistically significant, is of no practical importance. It would appear that the steps taken to ensure that subjects are motivated and highly trained had a great affect upon between subject variance. Comparison of the ω^2 of tables 5 and 10 illustrates this clearly.

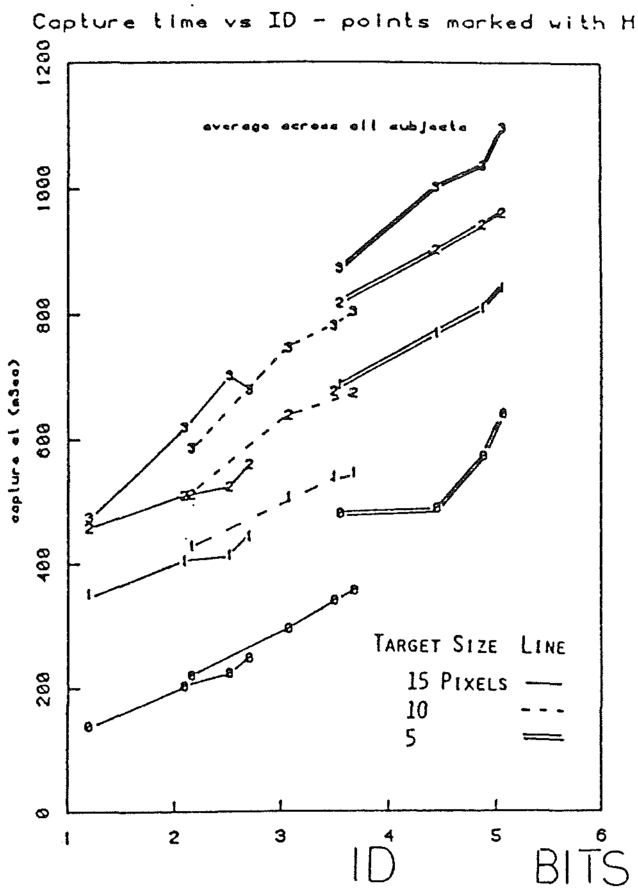

Capture time vs ID - points marked with H
average across all subjects
capture el (mSec)
TARGET SIZE LINE
15 PIXELS
10
5
ID
BITS

fig 6

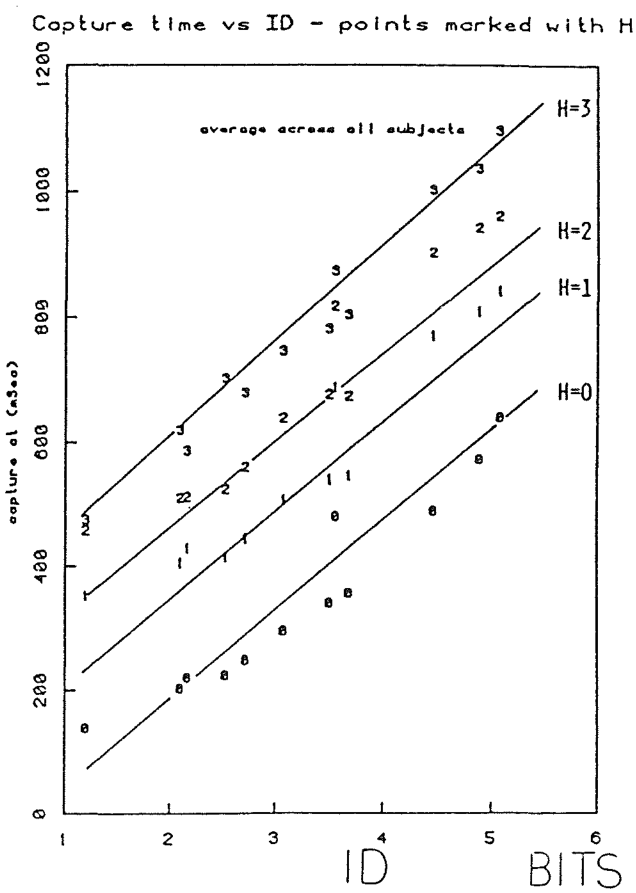

Capture time vs ID - points marked with H
average across all subjects
capture el (mSec)
H=3
H=2
H=1
H=0
ID
BITS

fig 7

Table 10 - Capture Time Anova - Contest

factor	F		MSE	p	ω^2
Subject	F(1,4)=	353	4945	< 0.001	-
ID	F(11,44)=	200	31	< 0.001	0.48
H	F(3,12)=	226	92	< 0.001	0.44
interact	F(33,132)=	2.0	17	0.03	0.004

Results of multiple regression analysis - Contest

Due to the greater number of cells in the experimental design of Contest, one is more inclined to have confidence in the results of regression analysis. Comparing tables 8 and 11 we see that the linearity of Contest data is less than that of the pilot experiment. Nevertheless, both the r^2 and RSE indicate a very good linear fit to the model of equation 6. It would appear that the combination of Fitts' Law and the Hick-Hyman Law stands up to the more realistic task of Contest almost as well as to the task of the pilot experiment.

The capture time data for Contest is presented in figure 6 with the multiple regression line drawn in for several values of H, and in figure 7 with equal size targets connected by lines. Examination of figure 7 shows that although the model explains some 95% of the variance, clearly target size plays a role besides the one recognized by ID. FIgure 7 shows how the two larger target sizes (about 0.5° and 1.0° of visual arc) fall in line whereas the smallest targets (about 0.3° of visual arc) seem to take longer to capture. Jagacinski and Monk [In Press], have tested Fitts Law in two dimensions using similar apparatus and found Fitts' Law to hold for targets of this size. Their criterion for target capture was not quite as simplistic, in that they allowed the cursor to leave the target for very short periods of time during the capture in order to "avoid penalizing the subjects for slight amounts of jitter" [Jagacinski and Monk, In Press]. It is possible that the stringent operational definition of capture used in Contest lengthened CTs for small targets by accentuating the effects of muscular tremour.

Table 11 - Capture Time regression analysis - Contest

for model: $CT=\alpha+\beta H+\gamma ID$

α	β	γ	r^2	RSE
-83 mSec	142 mSec/bit	144 mSec/bit	0.95	55

Practice effects

Practice effects were investigated by performing the analysis described above for each run of both the pilot study and Contest. Figure 8 is essentially the same graph that was displayed near the apparatus. It shows the anticipated flattening out of performance. Figure 9 shows how the form of the data does not change qualitatively from session to

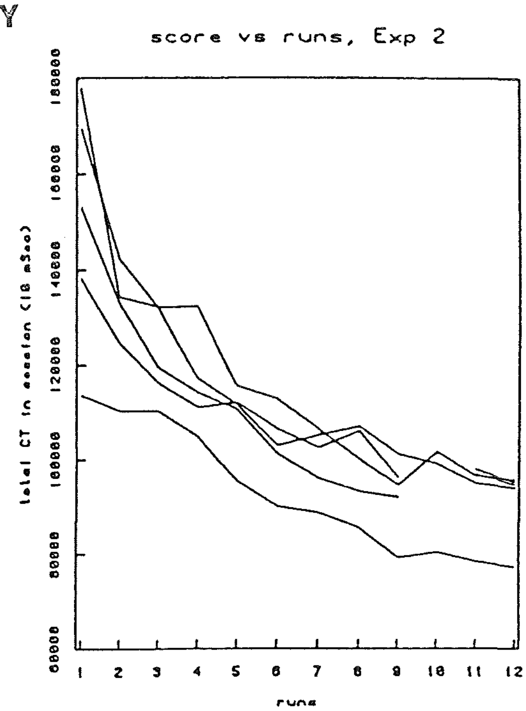

score vs runs, Exp 2
total CT in session (10 mSec)
runs

fig 8

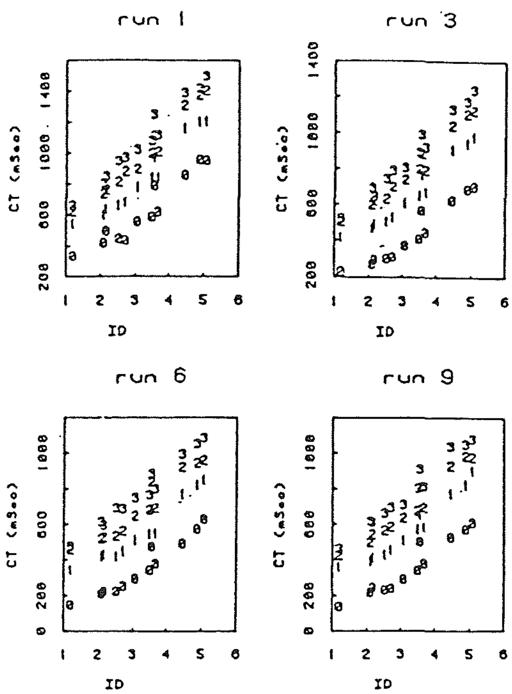

run 1
run 3
run 6
run 9
CT (mSec)
ID

fig 9

session even though quantitatively it can be seen to shift downwards.

Table 12
Regression analysis of Practice Effects - Contest

$T = \alpha + \beta H + \gamma ID$

run	α	β	γ	r^2	RSE
1	6.18	154	200	0.95	6.8
2	-47.7	149	184	0.95	6.5
3	-61.1	148	165	0.95	6.1
4	-76.5	140	158	0.95	5.8
5	-95.6	141	152	0.95	5.8
6	-75.9	141	143	0.95	5.6
7,8	NA	NA	NA	NA	NA
9	-90.6	141	149	0.95	5.6
10	-102	140	155	0.94	6.4
11	-68.8	141	143	0.94	5.8
12	-83.2	142	144	0.95	5.5
	mSec	mSec/bit	mSec/bit		

Perhaps the most interesting practice effect is evident in the CT data of the pilot experiment. Table 12 lists the regression results of each session of the pilot study whereas figure 10 shows the grand mean of RT, MT and CT for each run, pooled across subjects. We see that although MT decreases from the first session to the last session by about 13% RT changes little. One might have expected that with practice both RT and MT would decrease. Furthermore, the same trend is visible in the regression of CT. The intercept term increases even though the mean of CT gets smaller with practice. It would appear that as they learn the task subjects invest time at the beginning of each response which they can regain during the movement phase. Finally we observe that much improvement took place in the last two sessions of the pilot study, a clear indication that the subjects' performance had not stabilized.

In Contest, the intercept of the CT multiple regression started slightly greater than zero and steadily decreased. One assumes that the lack of temporal uncertainty in the task would chop a constant time out of the CT, but a negative intercept seems unrealistic at first glance. On closer examination one learns that several studies of discrete target capture behaviour found a negative Fitts' Law intercept [Fitts and Peterson, 1964]. The intercept of the regression in the extrapolation of the data to a point at which H and ID equal zero. Zero response entropy corresponds to the situation where the subject has no choice to make, and so is well defined. Zero Index of difficulty corresponds to an odd geometry in which the width of the target and the length of the motion are equal, an unrealistic scenario.

The regression coefficients of the CT regression decrease steadily with practice.

They vary inversely with information capacity and so we see the subjects' capacity to transmit information increasing with training.

Discussions and Conclusions

We have shown that CT is influenced by the degree of choice and by the required movement precision. By using a simple combination of Fitts' and the Hick-Hyman Laws, most of the variation in the data can be accounted for. These results hold both in a highly discrete and a continuous setting.

In the pilot experiment, RT was as well described by the Hick-Hyman Law as MT was by Fitts' Law. CT was as well described by the combination ("Hitts") law. The least squares multiple linear regression fit to the most (though still not fully) practiced data of the pilot experiment was:

$$CT = 260 + 130H + 150ID \quad (mSec)$$

whereas the most highly trained session of Contest yielded:

$$CT = -83 + 140H + 140ID \quad (mSec)$$

Following Fitts and Peterson [1964], the ID suggested by Welford, namely:

$$ID' = log_2(A/2r + 0.5) \qquad (7)$$

was tried out to see what effect it would have. The quality of fit was unchanged and the information capacity with respect to ID was reduced by about 11%. There seems no reason with these data to favour equation (7) over equation (2).

At risk of becoming embroiled in controversy, we comment that the data of the pilot experiment support the hypothesis that response and movement are executed sequentially as two separate stages. The independent measurement of RT and MT suggests that the factors affecting RT do not affect MT and vice versa. It would appear that sequential behaviour is a fact. Whether or not internal processing has the same structure is another problem altogether.

In the second experiment, the ability to independently measure RT and MT has been lost and so no such claim can be made. All that can be said here is that the factors which affected RT and MT separately in the pilot study interact to a negligible extent. One could probably analyze the time series of joystick positions with more sophisticated analysis and divide CT in the more continuous task Our experience has shown that that this may be difficult to accomplish.

Table 13 compares several studies in the literature with our results. We point out that although the experimental methods differ greatly (for instance the studies shown differ widely in modality of stimulus and response) information capacity with respect to H

varys over all by only about 25%. Fitts Law values, however, differ widely between experiments.

Table 13 - Comparison of results with other studies

Study	with respect to	
	ID	H
Hick[1952]	-	6.4
Hyman[1953]	-	7.87
Fitts[1954]	11.5	-
Jacinski and Monk[In Press]	5.0	-
Fitts and Peterson[1964]	13.5	-
Experiment 1	6.7	7.9
Experiment 2	6.9	7.0
	bps	bps

This would suggest to us that there is much to be gained at the physical interface level to the user. There is a need to tune the dynamics of mouse driven systems. This implies some quantitative method is required to provide a criterion for optimising design. To illustrate the environment in which software engineers typically work we quote from the notes for software developers included with what is one of the world's leading mice:

> We strongly urge you to try 2X magnification. Most software engineers are reluctant to do so, but after trying it, they find the feeling of control and speed far outweigh the inability to choose single pixels.. [p 8, Mouse Systems Corporation, M-2 Optical Mouse Technical Reference Manual, Jan. 1984]

Information capacity with respect to ID is a good starting point for tuning an interface. Anyone who has used a mouse recognizes that there are tasks which require higher or lower gains depending on the average size of targets and lengths of motions. One hopes that eventually a body of knowledge and guide lines will appear for what dynamics to use in which typical situations.

BIBLIOGRAPHY

Beggs, W.D.A., J.C. Graham, T.H. Monk, M.R.W. Shaw, and C.I. Howarth, "Can Hick's Law and Fitts' Law be combined?," *Acta Pschologica,* vol. 36, pp. 348-357, 1972.

Card, S. K., W. K. English, and B. J. Burr, "Evaluation of mouse, rate controlled joystick, step keys and Text keys for Text Selection on a CRT," *Ergonomics* , vol.21, pp. 601-613, 1978.

Fitts, P.M., "The Information capacity of the human motor system in controlling the amplitude of movement," *Journal of Experimental Psychology*, vol. 47, pp. 381-391, 1954.

Fitts, P.M. and J.R. Peterson, "Information capacity of discrete motor responses," *Journal of E xperimental Psychology,* vol. 67, pp. 103-112, 1964.

Hick, W.E., "On the rate of gain of information," *Quarterly Journal of Experimental Psychology*, vol. 4, pp. 11-26, 1952.

Hyman, R, "Stimulus Information as a Determinant of Reac tion Time,"*Journal of Experimental Psychology*, vol. 45, pp. 188-199, 1953.

Jagacinski, R. and D. L. Monk, "Fitts Law in two dimensions with hand and head movements,"*Journal of Motor Behaviour,* (In Press).

Kornblum, S., "Choice Reaction Time for Repetitions and Non-repetitions: A Rexamination of the Information Hypothesis," in *Attention and Performance* , ed. A. F. Sanders, pp. 178-187, North Holland Publishing Co., Amsterdam, 1967.

Meyers, B. A., "The User Interface for sapphire: A Screen Allocation Package Providing Helpful Icons and Rectangular Environments," InPress.

Sternberg, S., "The Discovery of Processing Stages: Extensions of Donders' Method.," in *Attention and Performance II,* ed. W. G. Koster, pp. 276-315, North Holland Publishing Co., Amsterdam, 1969.

Taylor, D.A., "Stage Analysis of Reaction Time,"*Psychological Bulletin,* vol. 83, no. 2, 1976.

Welford, A.T., *Fundamentals Of Skill*, Methuen and Co.Ltd, London, 1968.

This work was partly supported by funds provided under NSERC grant A1794 and NASA grant NAGW-429.