ISO9241 – Part 9

I. Scott MacKenzie
York University

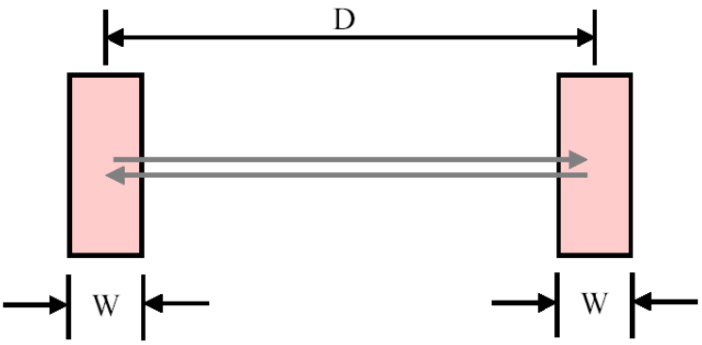
What is ISO 9241-9

- ISO 9241
  - "Ergonomic design of visual display terminals (VDTs) used for office work"
- Seventeen parts
  - Part 9: "Requirements for non-keyboard input devices"
  - Part 9 - Annex B – "Performance testing"

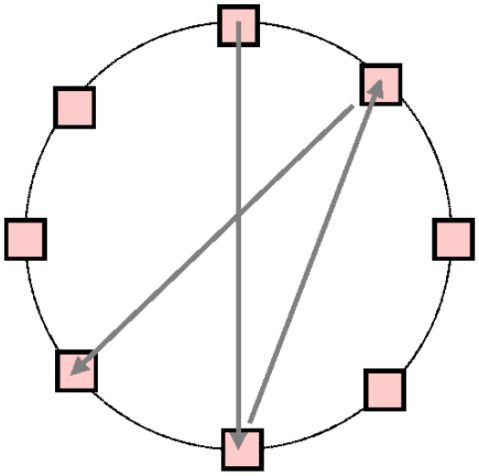
Annex B – Performance Tests

- One-directional tapping test
- Multi-directional tapping test
- Dragging test
- Path following test
- Tracing test
- Free-hand test
- Grasp and park (homing) test

One-directional Tapping Test



Multi-Directional Tapping Test

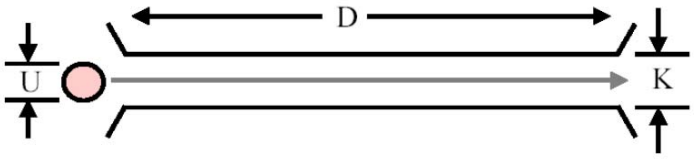


Note:
24 boxes

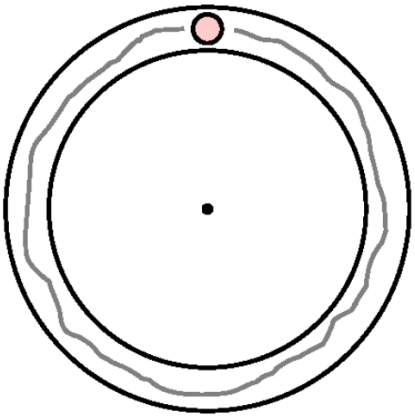
Dragging Test

- Variation on
  - One-directional tapping test
  - Multi-directional tapping test
- An object is “dragged” and dropped in the destination target
- See MacKenzie, Sellen, & Buxton paper from *CHI 1991*

Path Following Test



Tracing Test



Free-hand Input Test

Grasp and Park Test

- "The subject performs a simple pointing task and operates a key on the keyboard between each pointing task with the same hand."

Device Assessment Questionnaire

- ISO 9241-9 also includes a comprehensive qualitative evaluation
- Uses 13 questions and a 5-point Likert response scale
- Assesses aspects of operation, fatigue, comfort, and overall usability

Next slide

Device Assessment

Please circle the x that is most appropriate
as an answer to the given comment.

1. The force required for actuation was
   x x x x x
   too low too high

2. Smoothness during operation was
   x x x x x
   very rough very smooth

3. The mental effort required for operation was
   x x x x x
   too low too high

4. The physical effort required for operation was
   x x x x x
   too low too high

5. Accurate pointing was

x	x	x	x	x
easy				difficult

6. Operation speed was

x	x	x	x	x
too fast				too slow

7. Finger fatigue:

x	x	x	x	x
none				very high

8. Wrist fatigue:

x	x	x	x	x
none				very high

9. Arm fatigue:

x	x	x	x	x
none				very high

13

10. Shoulder fatigue:

x	x	x	x	x
none				very high

11. Neck fatigue:

x	x	x	x	x
none				very high

12. General comfort:

x	x	x	x	x
very uncomfortable				very comfortable

13. Overall, the input device was

x	x	x	x	x
very difficult to use				very easy to use

14

ISO Metric for Performance Testing

Throughput

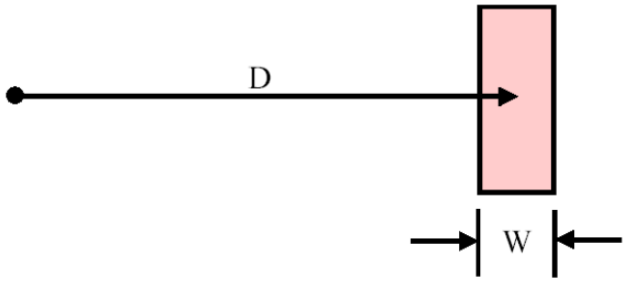
Throughput (1)

- Combines speed and accuracy in a single measure
- Based on Fitts' law (Fitts, 1954)
- Endorsed by ISO as the metric for evaluating pointing devices
- Calculations and procedures are tricky

Throughput (2)

- Benefits are substantial
- Not widely used in industry… yet!
- Aka "index of performance" or "bandwidth"
- We'll explain shortly, but first…

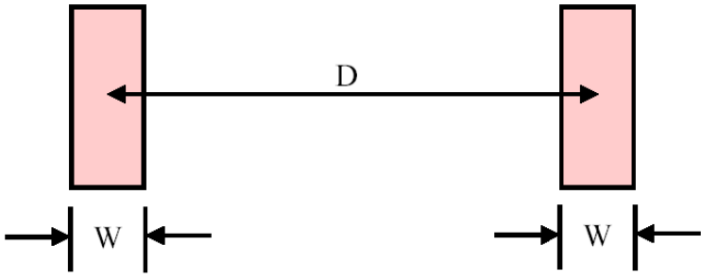
Fitts' Index of Difficulty (1)



$$ID = \log_2\left(\frac{D}{W} + 1\right)$$

Units: 'bits'

Fitts' Index of Difficulty (2)

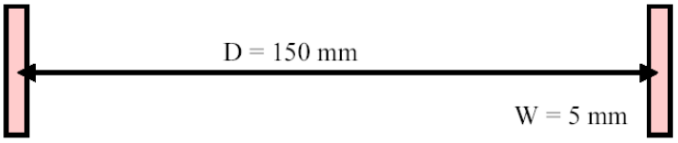


$$ID = \log_2\left(\frac{D}{W} + 1\right)$$

Units: 'bits'

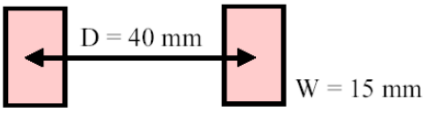
ID Examples

'Hard'



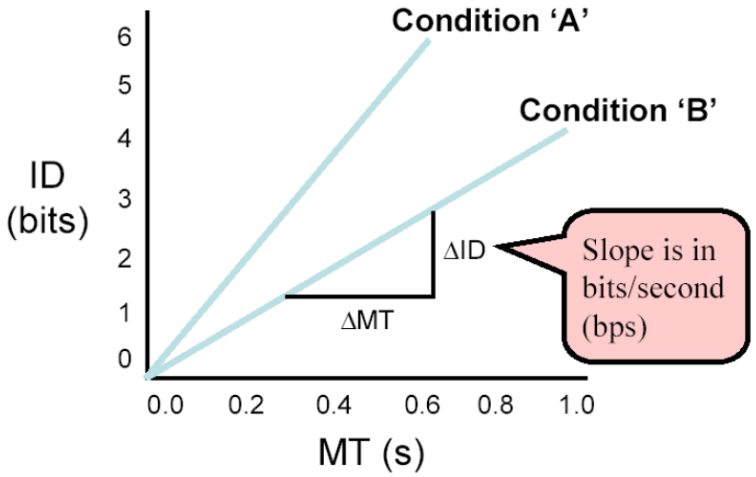
$$ID = \log_2\left(\frac{150}{5} + 1\right) = 4.9 \text{ bits}$$

'Easy'



$$ID = \log_2\left(\frac{40}{15} + 1\right) = 1.9 \text{ bits}$$

ID vs. Movement Time



Throughput

$$\text{Throughput} = \frac{ID}{MT} \text{ bits/s}$$

Some Data

D	W	ID (bits)	MT (s)
8	8	1.00	0.576
16	8	1.58	0.694
16	2	3.17	1.104
32	2	4.09	1.392
32	1	5.04	1.711
64	1	6.02	2.295
	Mean	3.48	1.306

$$\text{Throughput} = \frac{ID}{MT} = \frac{3.48}{1.306} = 2.66 \text{ bits/s}$$

Something is Missing

- The measure (2.66 bits/s) does not reflect the accuracy of participants' responses
- This seriously weakens within-study and between-study comparisons
- A solution was proposed by Crossman (1956) and endorsed by Fitts (1964), Welford (1968), and others

The Solution

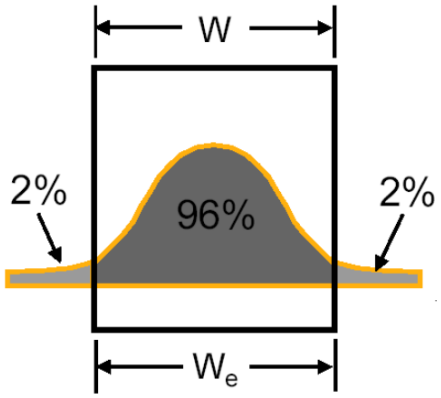
- Transform W into W_e
- W_e is the "effective target width"
- W_e = width of the distributions of "hits"
- W reflects what participants were asked to do
- W_e reflects what participants actually did

Effective Target Width

$$W_e = 4.133\ SD_x$$

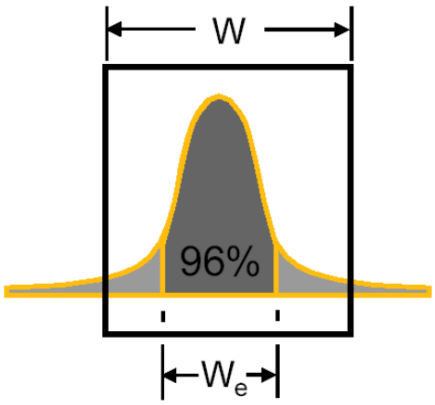
- SD_x – standard deviation in selection coordinates along axis of approach
- W_e is the central 96% of the spatial distribution
- Assumption: distribution is normal (research shows that the assumption holds reasonably well)
- Nominal error rate is 4%
- $H = \sqrt{2\pi e}\,\sigma = 4.133\ \sigma$ (Shannon, 1949)
- See MacKenzie (1992a, 1992b)

4% Errors



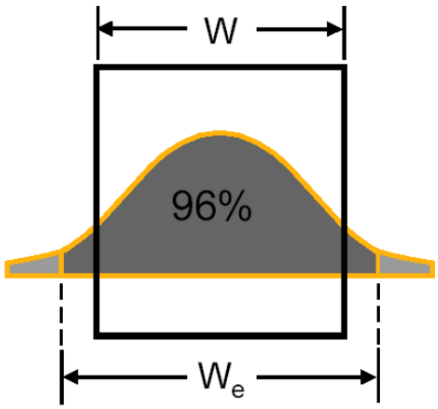
$W_e = W$

Less Than 4% Errors



$W_e < W$

Greater Than 4% Errors



$W_e > W$

"Effective" Index of Difficulty

$$ID_e = \log_2(D / W_e + 1)$$

- *ID* is the difficulty of the task presented to the participants
- ID_e is the difficulty of the task achieved by the participants
- ID_e reflects what participants actually did
- May use D_e instead of D, if participants did not "go the distance" (check data to determine this)

Some Data (Revisited)

D	W	W_e	ID_e (bits)	MT (s)
8	8	7.65	1.03	0.576
16	8	7.69	1.62	0.694
16	2	3.42	2.51	1.104
32	2	5.39	2.79	1.455
32	1	4.85	2.93	1.711
64	1	7.44	3.26	2.295
		Mean	2.36	1.306

$$\text{Throughput} = \frac{ID_e}{MT} = \frac{2.36}{1.306} = 1.81 \text{ bits/s}$$

Compare

- Speed only:
  - Throughput = 2.66 bits/s
- Speed and accuracy:
  - Throughput = 1.81 bits
- Difference:
  - +47%

So…

- We know how to calculate Throughput
- Where do we begin?

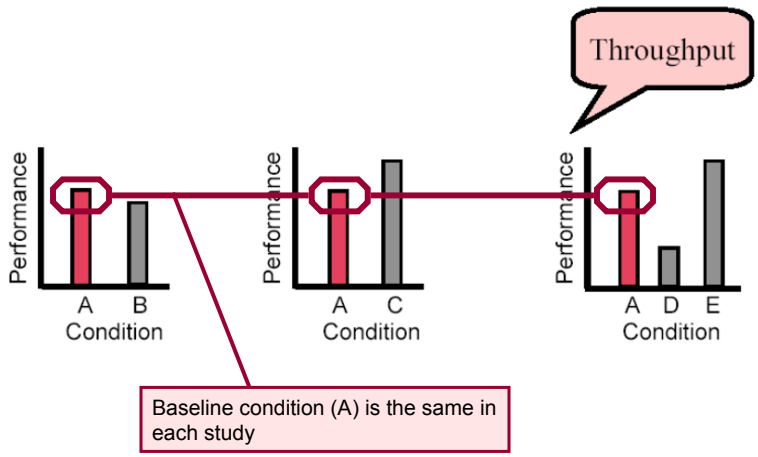
What is the Throughput for the Mouse?

- The question above is meaningless unless the test conditions are specified
- Conditions
  - Generic mouse
  - Standard mouse driver
  - Nominal gain setting
  - Serial task
  - Expert participants
  - Etc.
- Answer: $TP_{MOUSE} \approx 4.5$ bits/s
- The above is a "baseline condition"

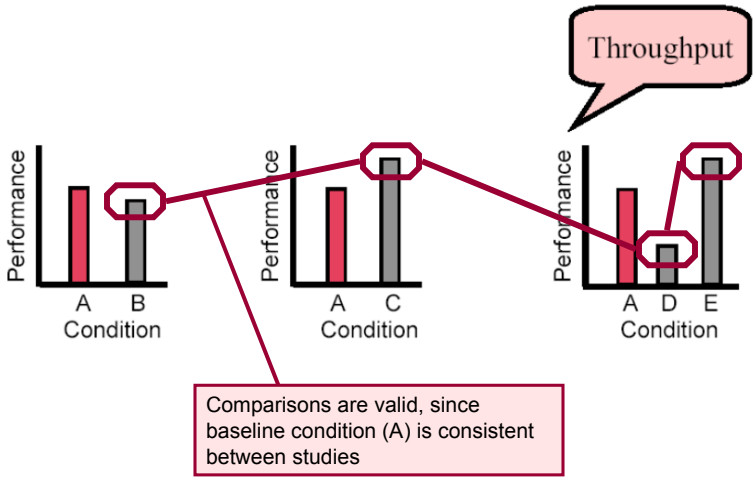
Baseline Conditions

- Serve as a check on experimental methods; i.e.,
  - Apparatus
  - Procedures
  - Data collection
  - Data analysis
- Obtaining a known and established result for a baseline condition means the methodology is sound
- Strengthens within-study comparisons
- Strengthens between-study comparisons

Clean up the Mess



Clean up the Mess



Example Study

- Four device conditions
  - Remote pointing
    - Gyration *GyroPoint* (air)
    - Interlink *RemotePoint*
  - Desktop pointing
    - Gyration *GryoPoint* (desk)
    - Microsoft *Mouse 2.0*

Reference:

MacKenzie, I. S., & Jusoh, S. (2001). An evaluation of two input devices for remote pointing. *Proceedings of the Eighth IFIP Working Conference on Engineering for Human-Computer Interaction - EHCI 2001*. pp. 235-249. Heidelberg, Germany: Springer-Verlag. (Click here to view)

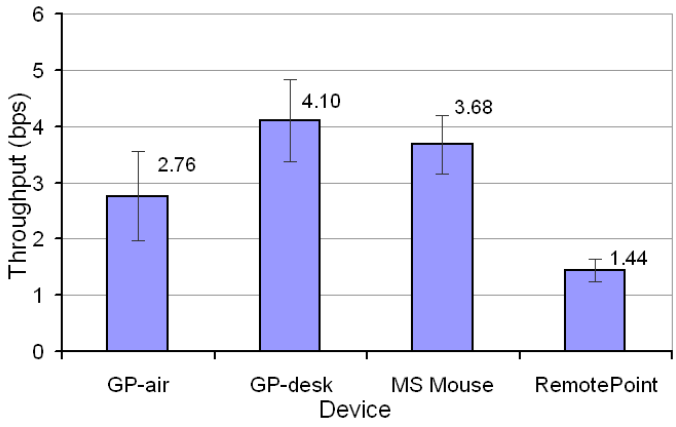
Gyration *GyroPoint*

39

Interlink *RemotePoint*

40

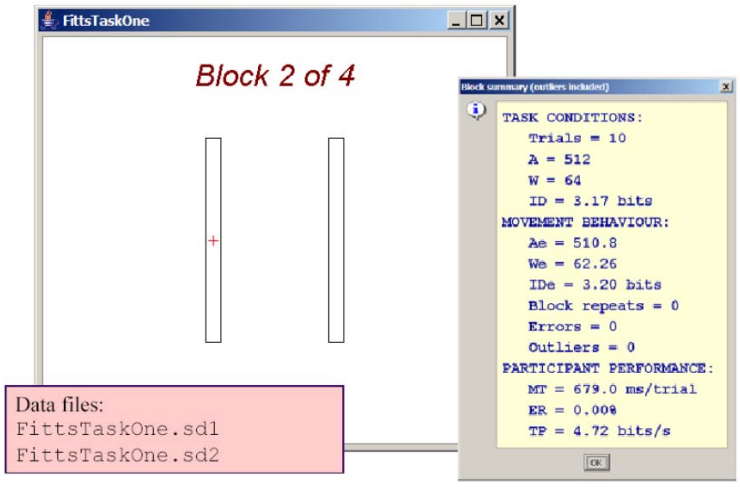
Results



Questionnaire – "Ease of Use" (Q13)

Participant	Device			
	Mouse 2.0	*GyroPoint*-desk	*GyroPoint*-air	*RemotePoint*
1	5	5	4	2
2	5	5	3	2
3	5	5	3	4
4	5	5	4	3
5	5	5	3	3
6	5	5	2	1
7	5	5	2	1
8	5	5	4	2
9	5	5	2	1
10	5	5	3	2
11	5	5	4	1
12	5	5	3	1
Mean	5.0	5.0	3.1	1.9
Note: 5 = easy, 4 = slightly easy, 3 = neutral, 2 = difficult, 1 = very difficult				

ISO 9241-9 Demo (Revisited)



Thank you

Questions?

Suggested Readings

1. Card, S. K., English, W. K., and Burr, B. J. Evaluation of mouse, rate-controlled isometric joystick, step keys, and text keys for text selection on a CRT, *Ergonomics 21* (1978), 601-613.
2. Carroll, J. M. (ed.), *Toward a multidisciplinary science of human-computer interaction*, (San Francisco: Morgan Kaufmann, 2003).
3. Kaindl, H. Methods and modeling: Fiction or useful reality?, *Extended Abstracts of CHI 2001*. (2001), 213-214.
4. Newell, A., and Card, S. K. The prospects for psychological science in human-computer interaction, *Human-Computer Interaction 1* (1985), 209-242.