



Human Reliability in Human Computer Interaction

Ned Hickling FIEHF

ned.hickling@synergy-ergs.com

IEHF
Registered
Consultancy

Acknowledgements & Caveats

- ONR for the original opportunity to do the original work (now extended)
- Views and opinions expressed are my own and do not necessarily reflect ONR views or policy

HCI and Error is Everywhere



What Am I Talking About?

Equipment Configuration Updated

Advantage Series 4100
User: SEC
Version 5.0.0 RC 1

Recipe: None

Overview Trends Login Emergency
Reports Alarms

Acid: pH Acid
Base: pH Base

pH Control

Max	0.00	0.00	gr
Actual	0.00	0.00	gr
	2.7 pH		
SP	1.0 pH		

1200.00 gr
Tare

Acid Base

N2

Stirrer Control

0.2 Nem
0 rpm
SP 0 rpm

Reactor Conditions

32.2 °C
2.7 pH
760 mBar
372.2 W/L
24,147.0 J
Act. SP 33.3 °C
End SP 55.0 °C

Data Logging Disabled
Log Comment

Auxiliary I/O

AI-1	0.0
AI-2	0.0
DI	#OFF
Soft-1	0.
Soft-2	0.
DO 1	On
DO 2	Off
AO	84

Aqueous

Target 0.00 gr
Actual 0.00 gr
0.00 gr/min
SP 0.00 gr/min

Feed 1 Control

1500.00 gr
Tare

Feed 2

Target 25.00 gr
Actual 0.00 gr
0.00 gr/min
SP 0.00 gr/min

Feed 2 Control

1500.00 gr
Tare

500 ml

Reactor Contents

Empty ±

Pressure Control

760 mBar
Act. SP 760 mBar
End SP 760 mBar

ACTIVE

Temperature Control

32.9 °C

TC Mode OFF
Jacket OFF
Reactor ON
Calibrate

Calorimetry Control

Temperature Control

Calorimetry Control

Overview of Paper Presentation

- Comparison of computer-based diagnostic data with THERP diagnostic error derivations
- Comparison of HCI holistic task data with object level data and the implications
- Comparison of THERP object level data with HCI data
- Conclusions for classical HRA methods application to HCI

The Problem

- UK safety cases are being submitted to UK Office for Nuclear Regulation to try and make a case where Human Computer Interaction is involved
- HRA methods are used intended for use on interfaces having discrete controls and displays,
- Is this tenable?

HRA Methods in Question

THERP	Has specific 'data' on diagnostic errors and on knobs and dials interaction errors
ASEP	Derived from THERP with no detailed interaction data, uses THERP diagnostic error derivation and is intended to be more conservative
SPAR-H	Assured by comparison to THERP: holistic judgement on interaction ergonomics quality and diagnostic judgements too

Method

- The comparison of published HCI error data and ‘corresponding’ existing HRA data for:
 - ★ Nuclear crew **diagnostic task** error probabilities
 - ★ **Holistic tasks**
 - ★ **Object level tasks** (e.g. WIMP)

The Comparison Problem

Classical Panel	Human Computer Interface
Walk to Panel	Select navigation object (x 2 or 3)
Select control	(On correct screen) select control icon
Operate control	Choose required action on pop-up menu
Confirm action effect as required	Interact with action object
	Confirm action effect as required
	[There may be intervening 'are you sure yes interactions too]

The Comparison Problem (cont'd)

- Design degrees of freedom are limited on classical panels and the number of interactions constrained
- Design degrees of freedom on HCI are much greater and the number of interactions required can vary far more

The Comparisons Process

- Direct functional equivalence cannot be achieved between knobs & dials with WIMPs
- Comparisons made between the **distributions of data sets**

Crew Diagnostic Error Probabilities

Scenario	Crew HEP	Ref.
Steam generator tube rupture 14 crews NOT using HCI	3E-02 (geometric mean)	[21]
Loss of coolant accident	1E-01	[22]
Straightforward loss of feedwater	2E-01	[23]
Complex loss of feedwater	9E-01	[24]
Varying fault combinations of: Steam generator tube rupture, steam line break, electrical bus failure, leaking or slow opening pressure relief valve, absent radiation indications.	1	[24]
	2E-01	[24]
	1.4E-01	[24]
	1.4E-01	[24]
	1.4E-01	[24]
	No errors	[24]
	No errors	[24]
Routine task: stopping all but one charging pump and closing isolation valves.	No errors	[24]

Derived HCI p(Diagnostic Error)

Missing data estimate	Resulting summary distributions		
	5th centile	Lognormal Mean	95th centile
'No error' data excluded	7E-02	2E-01	5E-01
3E-02 [21]	2E-02	1E-01	4E-01
1E-03	1E-02	5E-02	3E-01
THERP diagnosis screening	ca 30 mins	20 mins	10 mins
THERP diagnosis nominal	ca 20 mins	10 mins	Not specified

Example Objects, Interactions & Experimental Issues

- Icons
- Menus
- Data fields
- Keys
- Keyboards
- Drag and Drop
- Etc.
- Age
- Additional information
- Complexity/simplification
- Language
- Time
- Knowledge
- Etc.

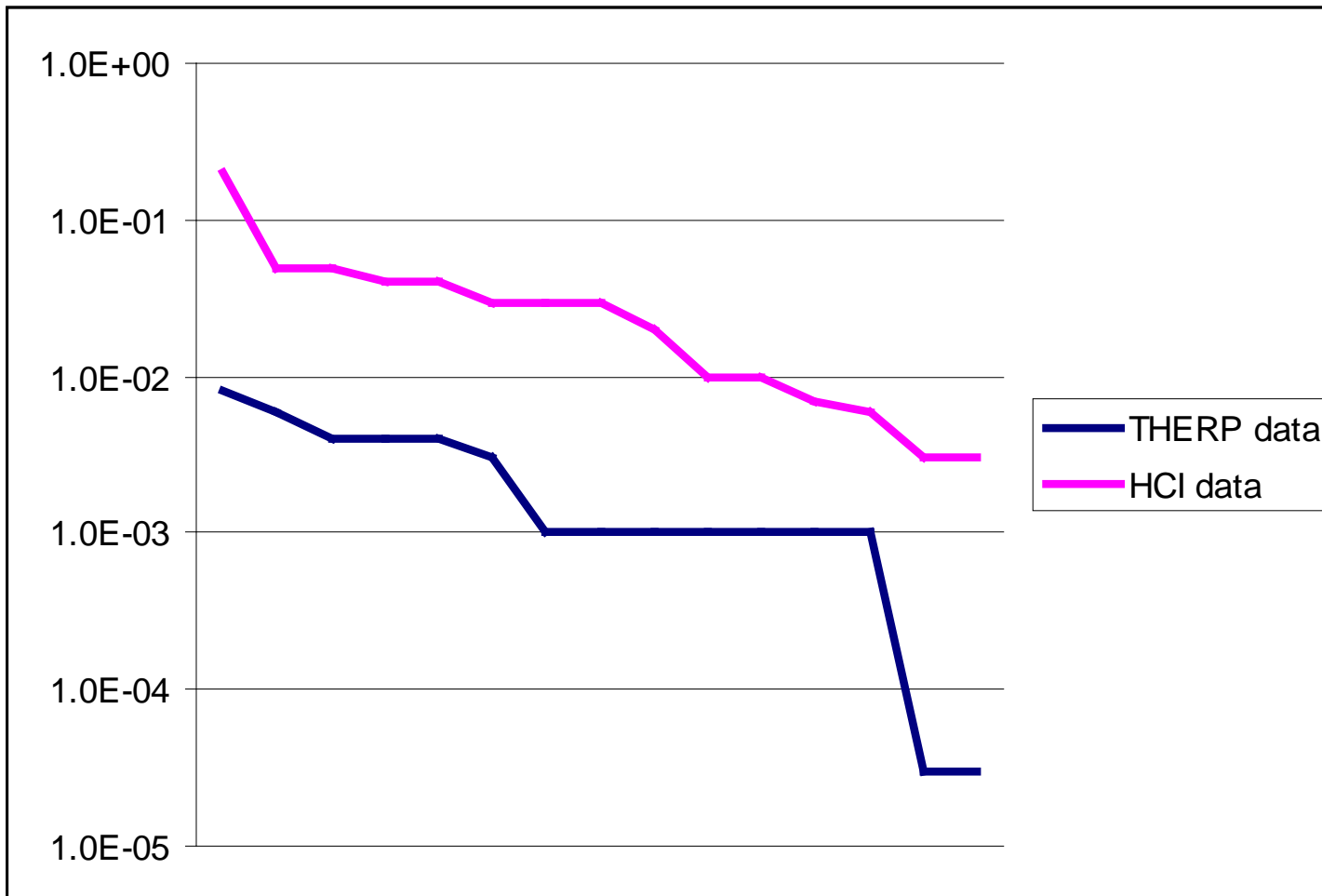
Object Level Data – THERP

- Definition: Interaction **on** a screen
- Selected examples of THERP object lvl data:
 - ★ 20-9,1 Selection of wrong display when it is dissimilar to adjacent displays: 3E-05
 - ★ 20-10, 3 Estimated HEPs for reading and recording quantitative information from chart recorder 8E-03
 - ★ 20-12,4 Select wrong control on a panel from an array of similar-appearing controls which are part of a well-defined mimic layout 1E-03
- 15 THERP data points selected with:
 - ★ No error producing conditions in descriptor
 - ★ Qualitative ‘equivalence’ to HCI object interactions
 - ★ Adjusted from median to mean estimate

Object Level Data – HCI Literature

- Examples:
 - ★ Icon selection with a double click
 - ★ Use of a shallow wide menu structure
 - ★ Process screen parameter reading with no parallax
- 15 HCI lit. object level data points selected:
 - ★ Judgement that error rate not reducible by further practice
 - ★ Confidently able to derive error data if not directly stated in the scrutinised paper
 - ★ No significant cognition (i.e. diagnosis or decision making) in the task
 - ★ **Most favourable HCI literature data put up for comparison to THERP**
 - ★ Data points *exclude* error recovery

HRA of Limited Application to HCI

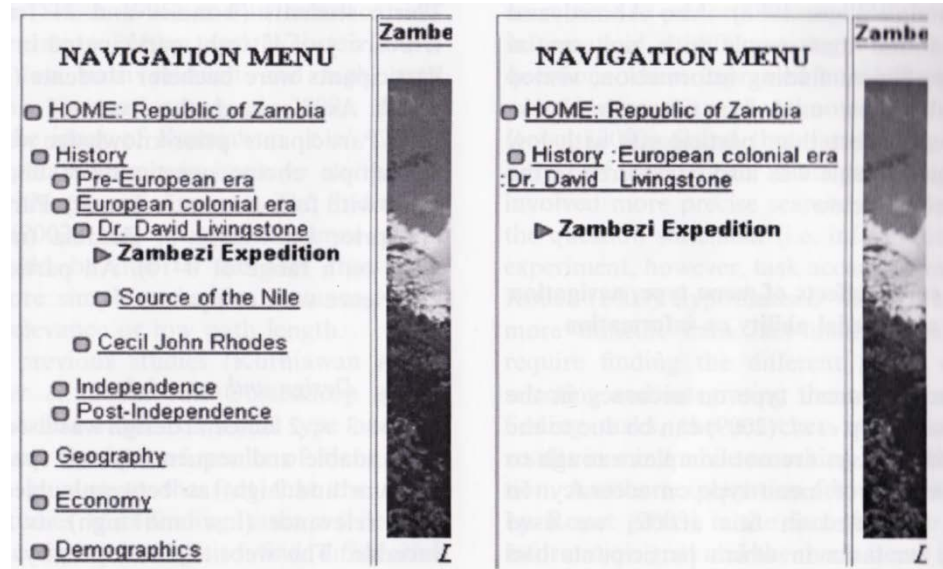


Significant HCI Error Producing Conditions

Condition	EPC	EF
Parallax	13	3.6
Long string data entry	12	3.5
Non-tactile keyboard	10	3.2
Infrequent automation experience & focus	8	2.8
Age	5	2.2
Interface language inconsistent with usage	4	2

Navigation

- Results from Melguizo et al. on navigation by expandable –v- sequential menus BIT 31, 1, pp 59-70
- Modify by HEART unfamiliarity (17)



HEP	Direct	Modified
Mean	0.23	0.01
SD	0.09	N/A
5 centile	0.1	0.006
95 centile	0.4	0.02

HCI Navigation conclusions

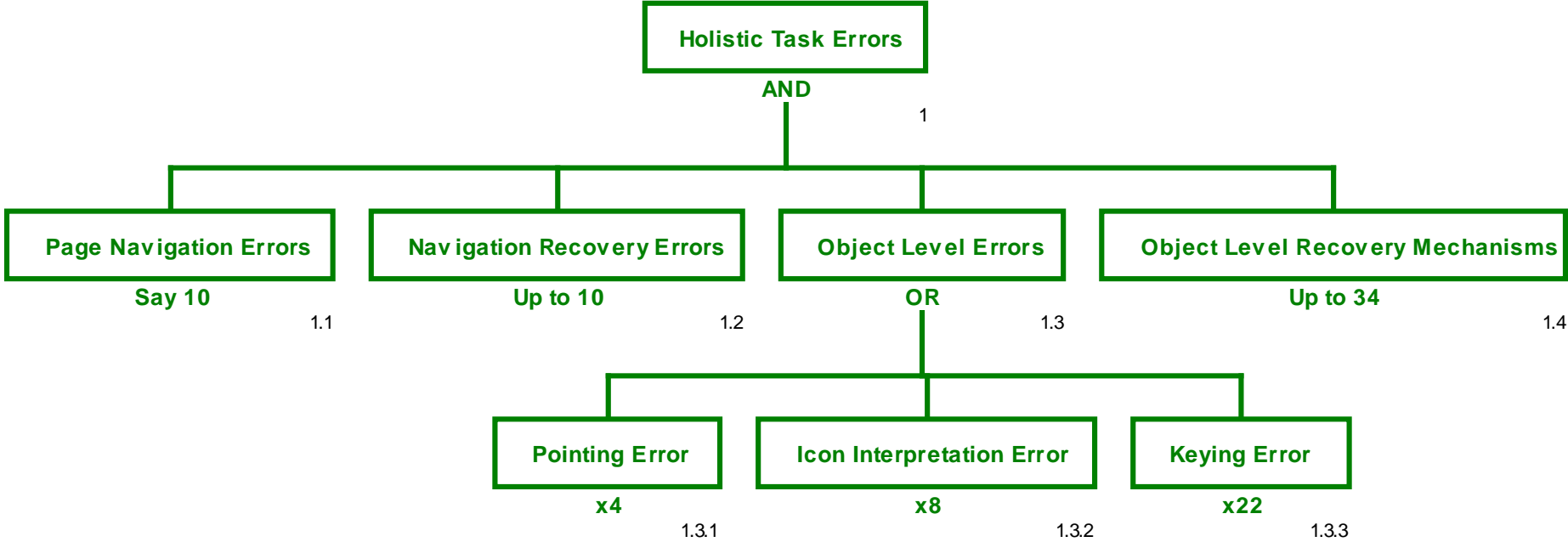
- The modification of Melguizo's data with EPC for unfamiliarity provides results entirely consistent with Snowberry on shallow wide –v- narrow deep menus (menu selection tasks)
- However navigation reliability cannot be dominated by menu structure – surely...
- Application specific experiments required

Holistic task data

➤ Sparse data:

- ★ 1st aid decision making, finite state automation knowledge, decisions on tabulated (alarm-like data),
Distributed team collaboration, NPP crew start up with automated procedures (simulation)

Holistic and Object Level Task Relationships



Object Level Data

-v-

Holistic HEP Data

	5 th centile	Lognormal mean	95 th centile
Object level	3E-03	2E-02	7E-02
Holistic	2E-03	1E-02	3E-01

Holistic- v- object level data conclusions

- It appears that holistic level tasks must succeed by containing persistent levels of self-recovery
- The only alternative would be that at least an order of magnitude improvement would be possible with further practice in object level tasks which seems unlikely given their inherent simplicity.

Implications for Classical HRA Methods

- In the UK (at least) HRA analysts appear to be calibrated to a THERPian HEARTian view of probability so ignorance of HCI data population characteristics will lead to optimistic judgements

High Integrity HCI Design

- Based on the foregoing we should assume that interactions will be unreliable
 - ★ Minimise initiating error potential
 - ★ Maximise error recovery opportunity

Navigation

- Initiation – new page selection
 - ★ NO numeric short cut keys (x 10)
 - ★ Shallow wide ungapped menu with hidden infrequently used items (x 3.2)
Labelled Icon with pop up help
- Recovery
 - ★ Similar destination pages distinguishable everywhere **at a glance without memory reliance**
e.g. background, watermark

Page Content

- Functionally label and provide contextual help for objects (x 8)
- Declutter by providing pop-up contextual help (x 4)
- Double click to select objects for interaction (x 4)
- Ensure task sequence is coherent , consistent and without conditional subroutines (x 3.25)

Interactions

- Data entry strings short at 3 digits (0.05, x 10)
- Double click to select objects for interaction (x 4)
- Avoid language complexity and non-preferred usage (as bad as mixed and foreign language) (x 5)

Overall Conclusions

- Reliance on HEART or THERP is likely to give optimistic results.
- However was THERP (in particular) optimistic anyway?
- HCI literature has data applicable to high integrity interfaces
- Don't be wowed just because it's sexy HCI: Interaction may not be reliable