

Formative Evaluation of Data Management Tools on Data Quality and Usability

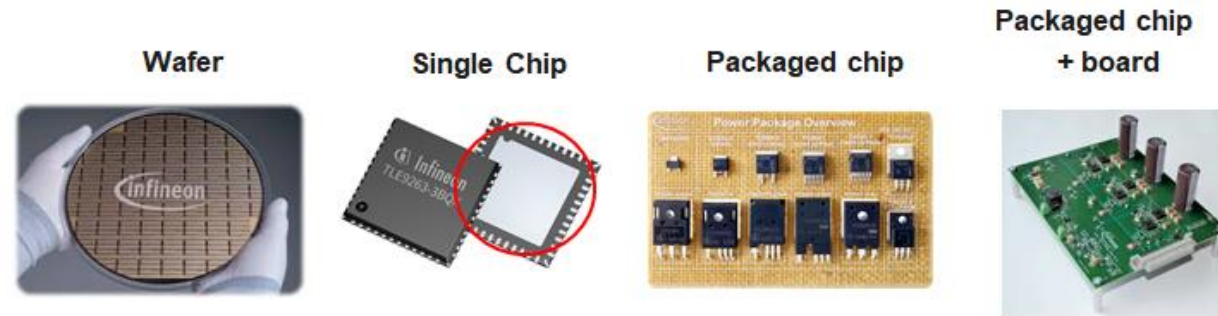
Supervisors: Dr.-Ing. Stefan Röhl, M.Sc. Bastian Busch (Infineon Technologies AG)

Hian Zing VOON

Bachelor's Thesis



- Better tool to manage test samples



 **Registration**

 **Planning**

 **Execution**

Shelf	Location Detail	Lot Number	Wafer Number	Technology	Responsible	Scrap Date
4	5	TU889614.03	1,2,3	HSTF1200	Max Turan	Dezember 24
4	5	TU223344	#3-#20	CIT110305	Franz Giebel	Dezember 25
4	4	TU911312	1,4,7,10,13,16,19	CIT120310	Kloppenburg	Dezember 30
4	4	QU778812.04	1,2,3	C9LMG	Kloppenburg	Dezember 30
4	4	QU778812.01	22,23,24,25	C9LMG_GG	Kloppenburg	Dezember 30
		PL887026	4	CARD 7	I.Kranz	Dezember 25
		P99667 UWU	1,2,3	INNOVAT TUM	Fromme	Dezember 25
4	3	GP0210201	13, 15, 17,18	P90TPA	Rebstock	Oktober 40
		GP2269622	9, 10	P90TPA	Rebstock	Oktober 40
4	3	HF151005.00	7	MOSFET - 10V	L.Balz	Dezember 25
4	3	HF888000.02	2	MOSFET - 10V	L.Balz	Dezember 25
4	3	PO118005.09	3	MOSFET - 10V	L.Balz	Dezember 25
		6ABB87HQZ2	4,5,6	CPTZ	H.Schaffer	
4	3	HF212007	18	CPTZ		Jan-25
1	5	QU106700	1,3,5,6,23,24	PMOS99R (GaN)	L.Deckert	Dec-25

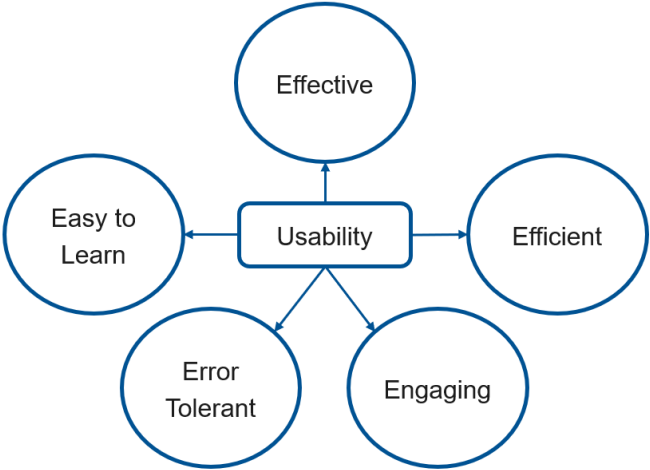
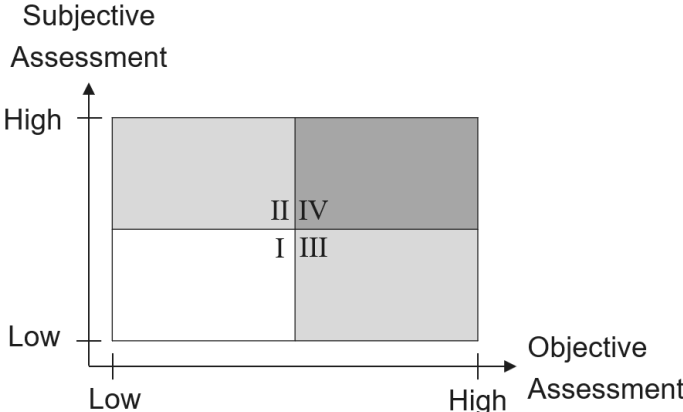
Background – Problem Statement

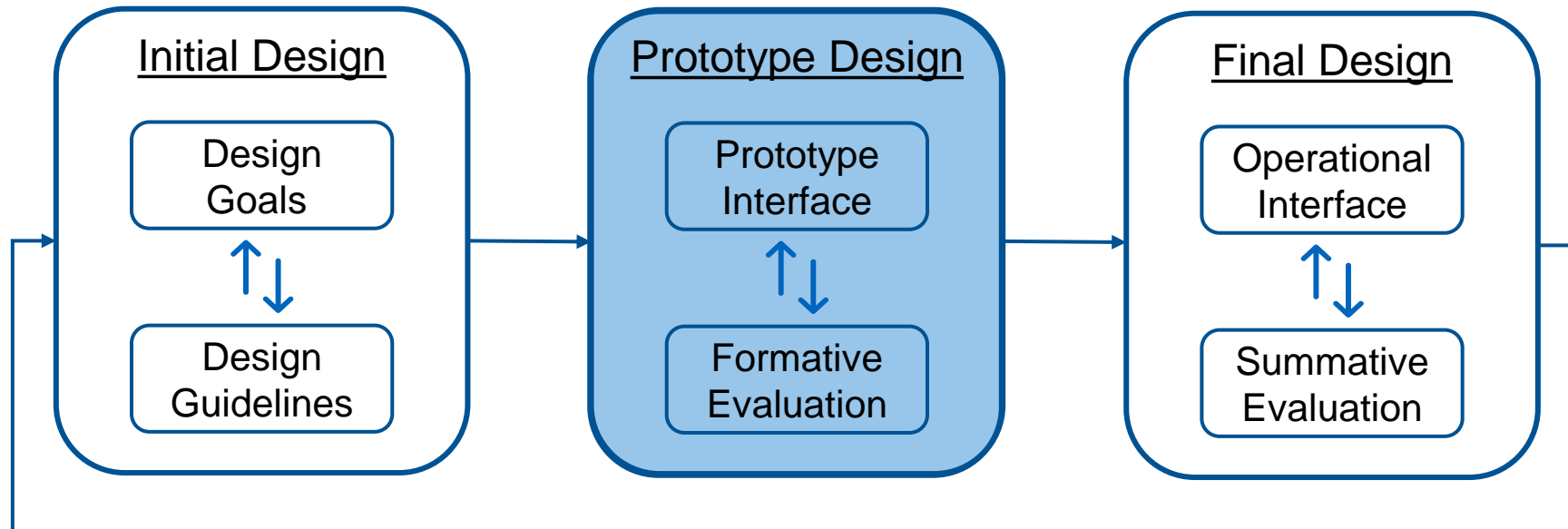


Shelf	Location Detail	Lot Number	Wafer Number	Technology	Responsible	Scrap Date
4	5	TU889614.03	1,2,3	HSTF1200	Max Turan	Dezember 24
4	5	TU223344	#3-#20 ←	Inconsistent data	Franz Giebel	Dezember 25
4	4	TU911312	1,4,7,10,13,16,19		Kloppenburg	Dezember 30
4	4	QU778812.04	1,2,3	C9LMG	Kloppenburg	Dezember 30
4	4	QU778812.01	22,23,24,25	C9LMG_GG	Kloppenburg	Dezember 30
		PL887026	4	CARD 7	I.Kranz	Dezember 25
		P9966		→ INNOVAT TUM	Fromme	Dezember 25
4	3	GP0210201	15, 15, 17, 18	P90TPA	Rebstock	Oktober 40
		GP2269622	9, 10	P90TPA	Rebstock	Oktober 40
4	3	HF151005.00	7	MOSFET - 10V	L.Balz	Dezember 25
4	3	HF888000.02	2	MOSFET - 10V	L.Balz	Dezember 25
4	3	PO118005.09	3	MOSFET - 10V	L.Balz	Dezember 25
		6ABB87HQZ2	4,5,6	CPTZ	→	Incomplete data
4	3	HF212007	18	CPTZ		Jan-25
1	5	QU106700	1,3,5,6,23,24	PMOS99R (GaN)	L.Deckert	Dec-25

Improve data quality

Improve system usability





- Data Quality
 - Accuracy
 - Completeness
 - Consistency

- Usability

Subjective Assessment

I found the information provided by the tool to be accurate.	Strongly disagree				Strongly agree
	1	2	3	4	5
I found the available data fields to be complete.	Strongly disagree				Strongly agree
	1	2	3	4	5
I found the data in the tool to be consistent.	Strongly disagree				Strongly agree
	1	2	3	4	5

I was satisfied with this tool overall.	Strongly disagree				Strongly agree
	1	2	3	4	5
I found this tool unnecessarily complex.	Strongly disagree				Strongly agree
	1	2	3	4	5
I found it easy to fill in the data.	Strongly disagree				Strongly agree
	1	2	3	4	5
I think that I would need the support of a technical person to be able to use this tool.	Strongly disagree				Strongly agree
	1	2	3	4	5
I had a clear understanding of where each data needs to be entered.	Strongly disagree				Strongly agree
	1	2	3	4	5
I found it easy to understand the correct format for entering the data.	Strongly disagree				Strongly agree
	1	2	3	4	5
I received support from the tool for entering the data.	Strongly disagree				Strongly agree
	1	2	3	4	5

Objective Assessment

$$CA_i = \frac{\sum_{j=1}^a acc(r_j, D(r_j))}{|a|}$$

$$CC_i = 1 - \frac{Y}{X}$$

$$CCS_i = 1 - \frac{Y}{X}$$

$$\sum_{i=1}^n CW_i \times CA_i ; \quad \sum_{i=1}^n CW_i \times CC_i ; \quad \sum_{i=1}^n CW_i \times CCS_i$$

Record time taken,
mouse clicks and
keystrokes

Implementation – Prototype Development



NEXTREL

Hian Zing Voon

243 samples

Status: Select...
Type: Select...
Responsible: -
Creator: -
Technology: -
Lot Num:
Basic Type: Select...
Wafer Num:
SpaRQ Project: Select...
Scrap Timeframe:

Reset Filter

	Sample Type	Lot Number	Wafer Number	Technology	Status
1	Wafer	TU180333	3	CIT001	In Storage
2	Wafer	TU180333	4	CIT001	In Storage
3	Wafer	TU180333	5	CIT001	In Storage
4	Wafer	TU180333	6	CIT001	In Storage
5	Wafer	TU180333	7	CIT001	In Storage
6	Wafer	TU180333	8	CIT001	In Storage
7	Wafer	TU180333	9	CIT001	In Storage
8	Wafer	TU285748	1	CIT002	In Storage
9	Wafer	TU285748	2	CIT002	In Storage
10	Wafer	TU285748	3	CIT002	In Storage
11	Wafer	TU285748	4	CIT002	In Storage
12	Wafer	TU285748	5	CIT002	In Storage
13	Wafer	TU285748	6	CIT002	In Storage
14	Wafer	TU285748	7	CIT002	In Storage
15	Wafer	TU285748	8	CIT002	In Storage
16	Wafer	TU285748	9	CIT002	In Storage
17	Wafer	TU285748	10	CIT002	In Storage

1 to 100 of 243 records

New sample

Select a sample...

Details

Sample Type: * Wafer
Description:
Lot Number: *
Project: No project
Technology: -
Status: * Scrap
Responsible: * Hian Zing Voon
Use Case: * Monitoring
Scrap Date Planned: Select a date
Scrap Date Actual: Select a date

Location

Has Location:
Site:
Building:

Save Changes

NEXTREL

243 samples

Status: Select...
Type: Select...
Responsible: -
Creator: -
Technology: -
Lot Num:
Wafer Num:
Project: Select...
Scrap Timeframe:

Reset Filter

Create new sample

Sample Type: * Select...
Sample Status: * Select...
Lot Number: * Enter your lot number
Wafer Number: *
Technology: Select...
Project: No project
Responsible: * Select...
Creator: * Select...
Use case: * Select...
Shelf: Select...
Location Detail: Select...

Hian Zing Voon

New sample

Select a sample...

Wafer
No project
In Storage
Hian Zing Voon
Select a sample to edit...
Qualification
Select a date
Select a date

CW_1
= 0.05

CW_2
= 0.05

CW_3
= 0.25

CW_4
= 0.20

CW_5
= 0.15

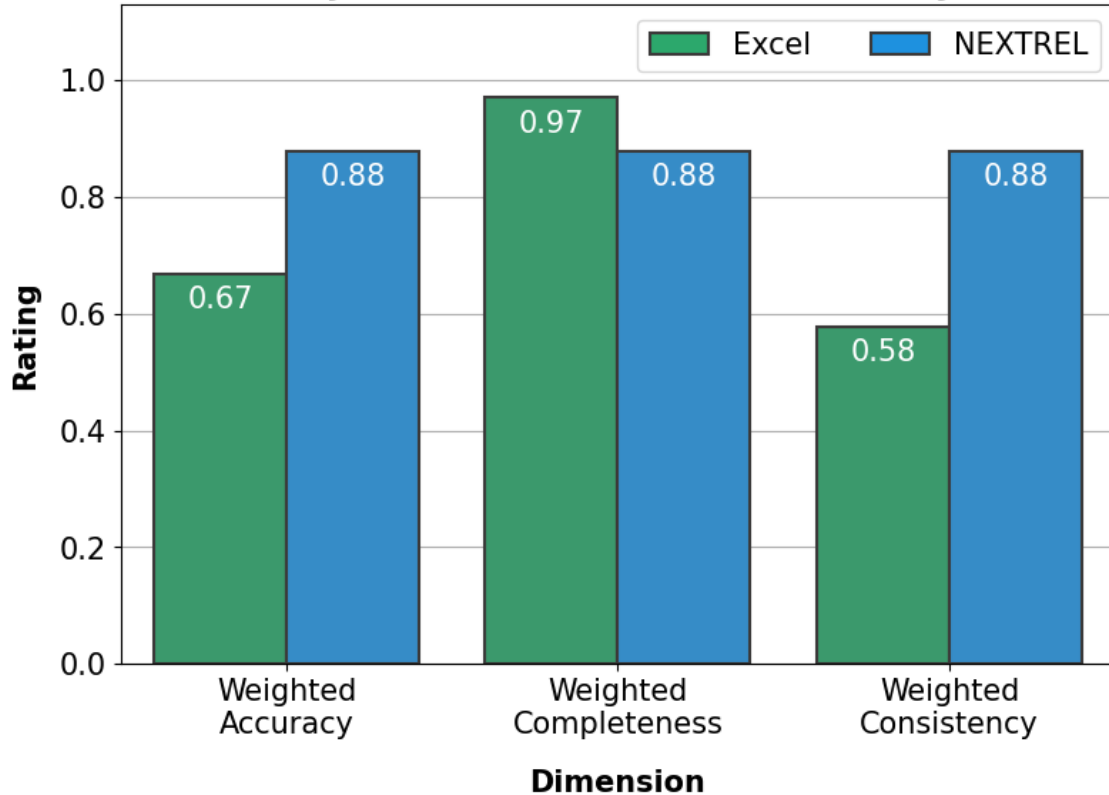
CW_6
= 0.20

CW_7
= 0.10

Shelf	Location Detail	Lot Number	Wafer Number	Technology	Responsible	Scrap Date
4	5	TU180333	3	CIT001	Linus Thorun	2029-12-01 00:00:00
4	5	TU180334	4	CIT001	Linus Thorun	2029-12-01 00:00:00
4	5	TU180335	5	CIT001	Linus Thorun	2029-12-01 00:00:00
4	5	TU180336	6	CIT001	Linus Thorun	2029-12-01 00:00:00
4	5	TU180337	7	CIT001	Linus Thorun	2029-12-01 00:00:00
4	5	TU180338	8	CIT001	Linus Thorun	2029-12-01 00:00:00
4	5	TU180339	9	CIT001	Linus Thorun	2029-12-01 00:00:00
-	-	-	1	KEY_8	-	2027-12-01 00:00:00
w2	3	QU664520	23	PPTX2	Max Turan	2027-12-01 00:00:00
w2	3	QU664522	9	PPTX3	Max Turan	2027-12-01 00:00:00
w2	3	QU664530	22	PPTX4	Max Turan	2027-12-01 00:00:00
w2	3	QU664550	15	PPTX5	Max Turan	2027-12-01 00:00:00
w2	3	QU664555	9	PPTX6	Max Turan	2027-12-01 00:00:00
3	2	HF888000.02	2	-	Leon Balz	2026-05-01 00:00:00
w2	2	GP226922.03	15	P90TPA	Hans Rebstock	-

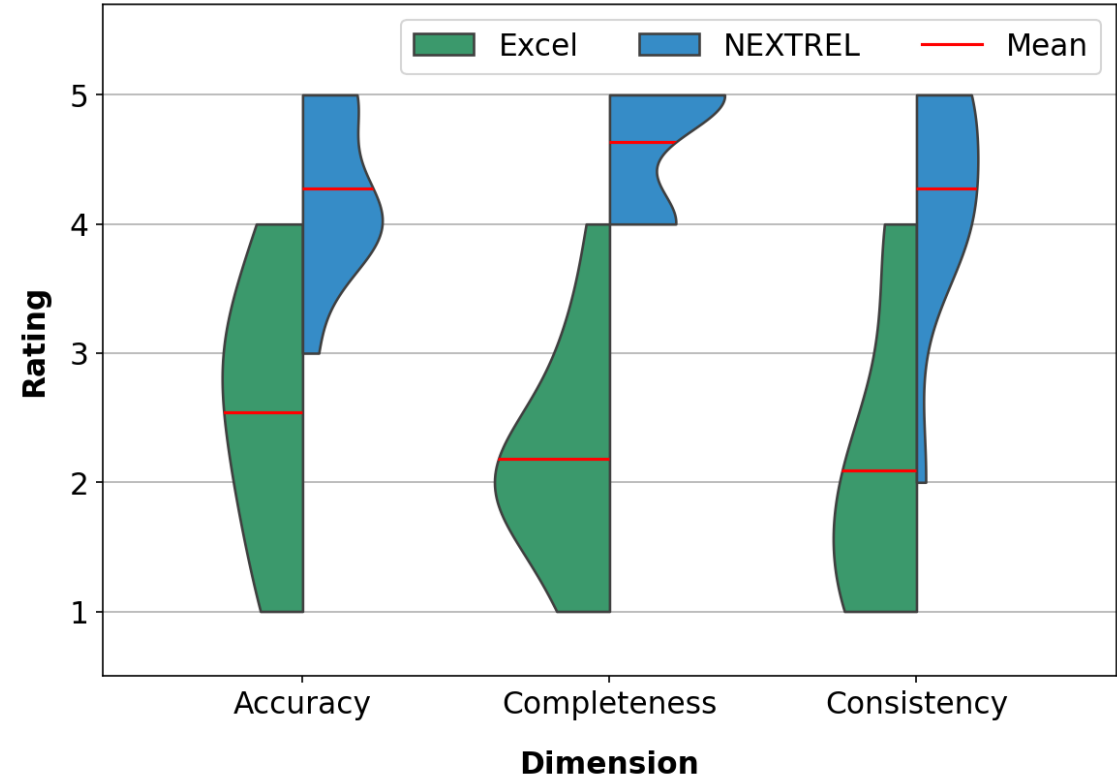
Objective Assessment

Objective Assessment of Data Quality

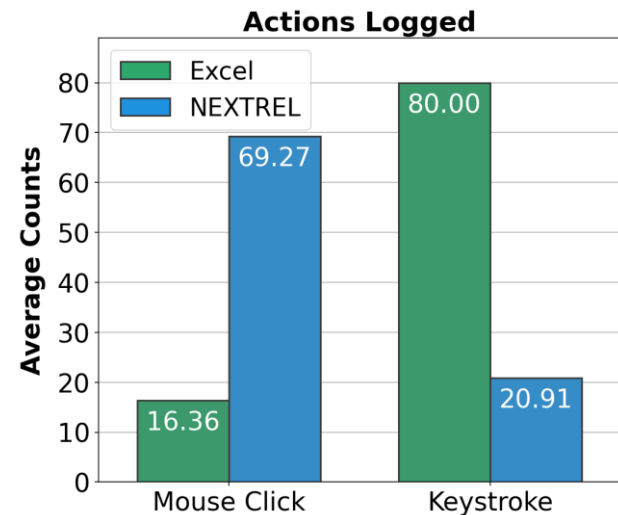
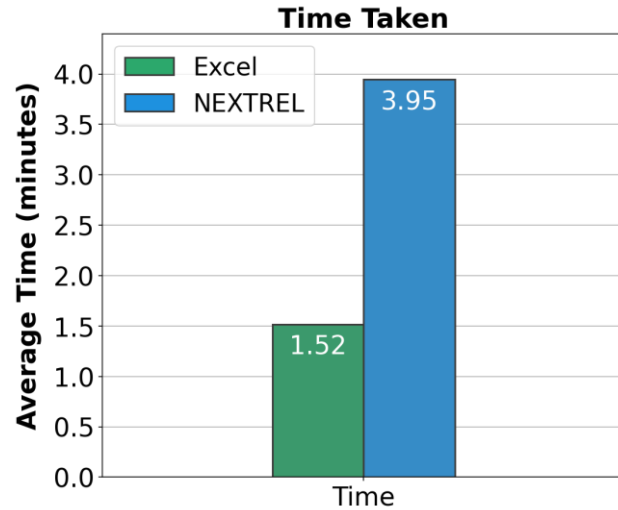


Subjective Assessment

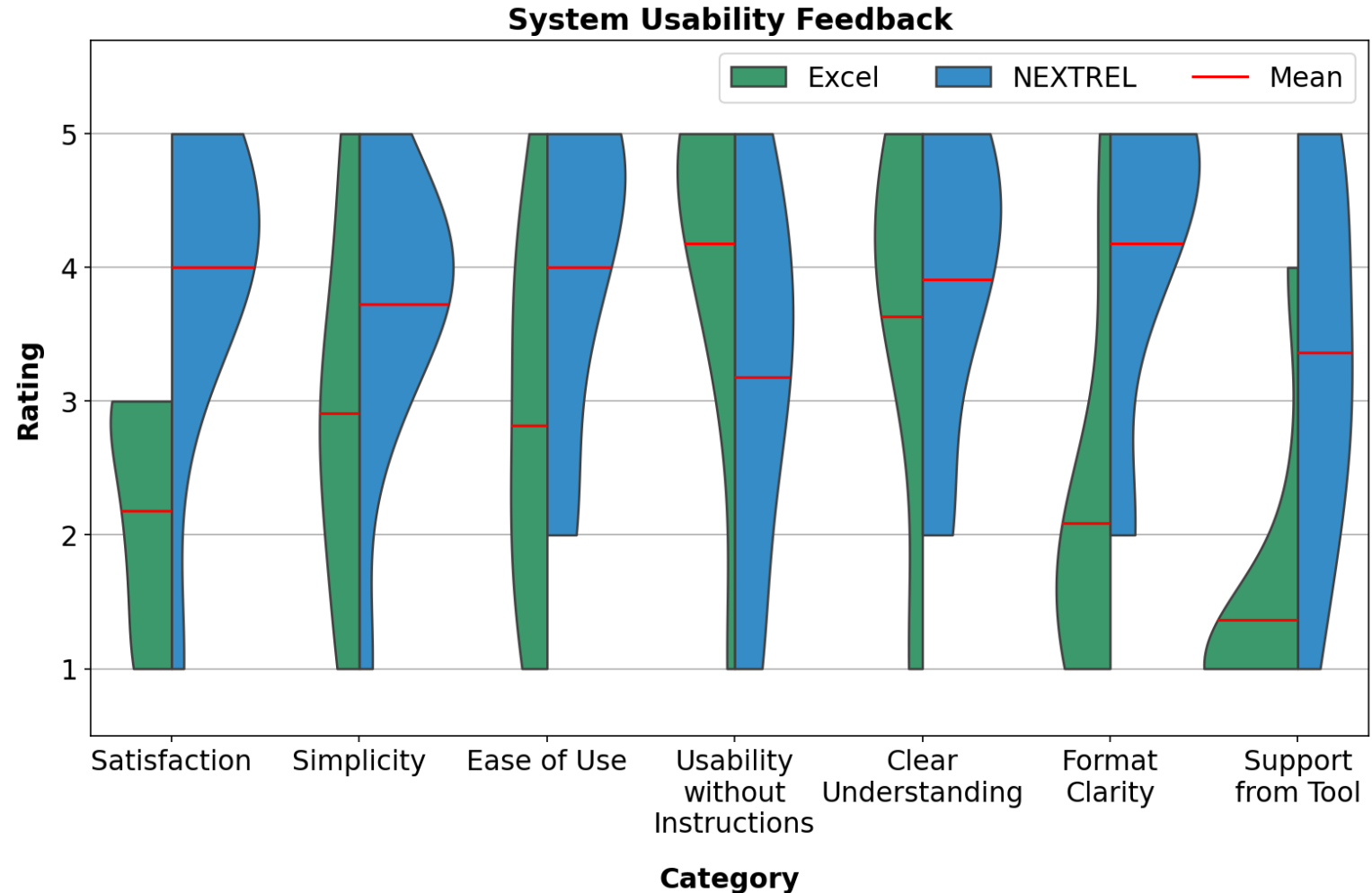
Subjective Assessment of Data Quality



Objective Assessment



Subjective Assessment



- Improved data quality and usability
- *NEXTREL* requires longer time and training for new users
- Standardized templates improve data quality

Thank you for your attention

- Accuracy
 - Syntactic accuracy
 - Semantic accuracy

Name	Address	Postal Code	City	Delivery Date
Alice	Arcisstraße 21	80333	München	01.09.2024
Bob	Boltzmannstraße 15	10623	Garching bei München	15.09.2024
Charlie	Am Campeon 1-15	85579	Neubiberg	30.09.2024

$$Acc(t) = \frac{\sum_{i=1}^t acc(r_i, D(r_i))}{|t|},$$

String errors $acc(r_i, D(r_i)) = \begin{cases} 1, & \text{if } r_i \in D(r_i), \\ 1 - NED(r_i, D(r_i)) & \text{otherwise.} \end{cases}$

Numerical errors $acc(r_i, D(r_i)) = \begin{cases} 1, & \text{if } r_i \in D(r_i), \\ 1 - \frac{|r_i - D(r_i)|}{Max(r_i, D(r_i))} & \text{otherwise.} \end{cases}$

Technology	
HSTF1200	M
CIT110305	F
CIT120310	K
C9LMG	K
C9LMG_GG	K
CARD 7	I.
INNOVAT TUM	F
P90TPA	R
P90TPA	R
MOSFET - 10V	L
MOSFET - 10V	L
MOSFET - 10V	L
CPTZ	†
CPTZ	
PMOS99R (GaN)	L

$$CA_i = \frac{\sum_{j=1}^a acc(r_j, D(r_j))}{|a|}$$

INNOVATIV_TUM

NED = Normalized Edit Distance (0,1)

Levenshtein Distance between
 “INNOVAT TUM” and “INNOVATIV_TUM”
 = 3

$$NED = \frac{3}{13}$$

$$acc = 1 - \frac{3}{13} = 0.77$$

Technology		
HSTF1200	1	M
CIT110305	1	F
CIT120310	1	K
C9LMG	1	K
C9LMG_GG	1	K
CARD 7	1	I.
INNOVAT TUM	0.77	
P90TPA	1	R
P90TPA	1	R
MOSFET - 10V	1	L
MOSFET - 10V	1	L
MOSFET - 10V	1	L
CPTZ	1	†
CPTZ	1	
PMOS99R (GaN)	1	

$$CA_{technology} = \frac{14.77}{15}$$

$$CA_{technology} = 0.9847$$

- **Completeness**

- Schema completeness
- Column completeness
- Population completeness

Person ID	Name	Gender	Birth Date	Email
1	Danny	Male	01.01.1990	danny90@gmail.com
2	Emily	Female	10.02.1993	NULL
3	Frank	Male	15.03.1995	NULL
4	Gabrielle	Female	20.05.1992	NULL

$$\text{Completeness rating} = 1 - \left(\frac{\text{Number of incomplete items}}{\text{Total number of items}} \right).$$

$$CC_i = 1 - \frac{Y}{X}$$

Scrap Date
Dezember 24
Dezember 25
Dezember 30
Dezember 30
Dezember 30
Dezember 25
Dezember 25
Oktober 40
Oktober 40
Dezember 25
Dezember 25
Dezember 25
Jan-25
Dec-25

$$CC_i = 1 - \frac{Y}{X}$$

Oct-24

Cell completeness
for this particular cell
= 0

	Scrap Date
1	Dezember 24
1	Dezember 25
1	Dezember 30
1	Dezember 30
1	Dezember 30
1	Dezember 25
1	Dezember 25
1	Oktober 40
1	Oktober 40
1	Dezember 25
1	Dezember 25
1	Dezember 25
1	
0	Jan-25
1	Dec-25

$$CC_{Scrap\ Date} = 1 - \frac{1}{15}$$

$$CC_{Scrap\ Date} = \frac{14}{15}$$

- Consistency

- Integrity constraint
 - intrarelation constraint
 - interrelation constraint
- Between 2 related data elements
- Format

Employee ID	Name	Age	Start Year	Position
100	Harry	22	2024	Reliability Engineer
101	Ivy	25	2023	Sales Representative
102	Jack	15	2022	IT Specialist

Table 2.3.: Example of an *Employee* relation

Promotion ID	Employee ID	Promotion Year	New Position
321	100	2022	Senior Reliability Engineer
322	101	2026	Sales Manager
323	102	in 5 years	Senior IT Specialist

Table 2.4.: Example of a *Promotion* relation for employees in Table 2.3

$$\text{Consistency rating} = 1 - \left(\frac{\text{Number of inconsistent units}}{\text{Total number of consistency checks performed}} \right)$$

$$CCS_i = 1 - \left(\frac{\text{Number of incomplete items}}{\text{Total number of items}} \right)$$

Wafer Number	
1,2,3	F
#3-#20	C
1,4,7,10,13,16,19	C
1,2,3	C
22,23,24,25	C
4	C
1,2,3	II
13, 15, 17,18	P
9, 10	P
7	N
2	N
3	N
4,5,6	C
18	C
1,3,5,6,23,24	P

$$CCS_i = 1 - \left(\frac{\text{Number of incomplete items}}{\text{Total number of items}} \right)$$

3

Cell consistency for this particular cell = 0

Wafer Number	
1,2,3	0
#3-#20	0
1,4,7,10,13,16,19	0
1,2,3	0
22,23,24,25	0
4	1
1,2,3	0
13, 15, 17,18	0
9, 10	0
7	1
2	1
3	1
4,5,6	0
18	1
1,3,5,6,23,24	0

$$CCS_{\text{Wafer Number}} = 1 - \frac{10}{15}$$

$$CCS_{\text{Wafer Number}} = \frac{5}{15}$$

The screenshot shows the NEXTEL software interface. A 'Create new sample' dialog is open, with a modal titled 'Edit Wafer Selections' in the foreground. The modal contains a 5x5 grid of wafer numbers (1-25) and two radio buttons: 'Select All' (unselected) and 'Selectable Range' (selected). The background shows a list of 243 samples with various filters and a 'New sample' button.

Create new sample

Sample Type: * Wafer

Sample Status: * In S

Lot Number: * TUS

Wafer Number: *

Technology: * Sel

Project: * No

Responsible: * Sel

Creator: * Sel

Use case: * Sel

Shelf: * Sel

Location Detail: * Select...

Save sample

Cancel

Edit Wafer Selections

Select All Selectable Range

1	2	3	4	5
6	7	8	9	10
11	12	13	14	15
16	17	18	19	20
21	22	23	24	25

Ok Cancel

NEXTEL

243 samples

Status: Select...
Type: Select...
Responsible: -
Creator: -
Technology: -
Lot Num:
Wafer Num:
Project: Select...
Scrap Timeframe:

Reset Filter

Hian Zing Voon

New sample

Select a sample...

Wafer
No project
In Storage
Hian Zing Voon
Select a sample to edit...
Qualification
Select a date
Select a date

Create new sample

Sample Type: * Wafer x v

Sample Status: * In Storage x v

Lot Number: * TU385354 **Validate Lot**

Wafer Number: * 1-25

Technology: Select... v

Project: No project v

Responsible: * Hian Zing Voon x v

Creator: * Hian Zing Voon x v

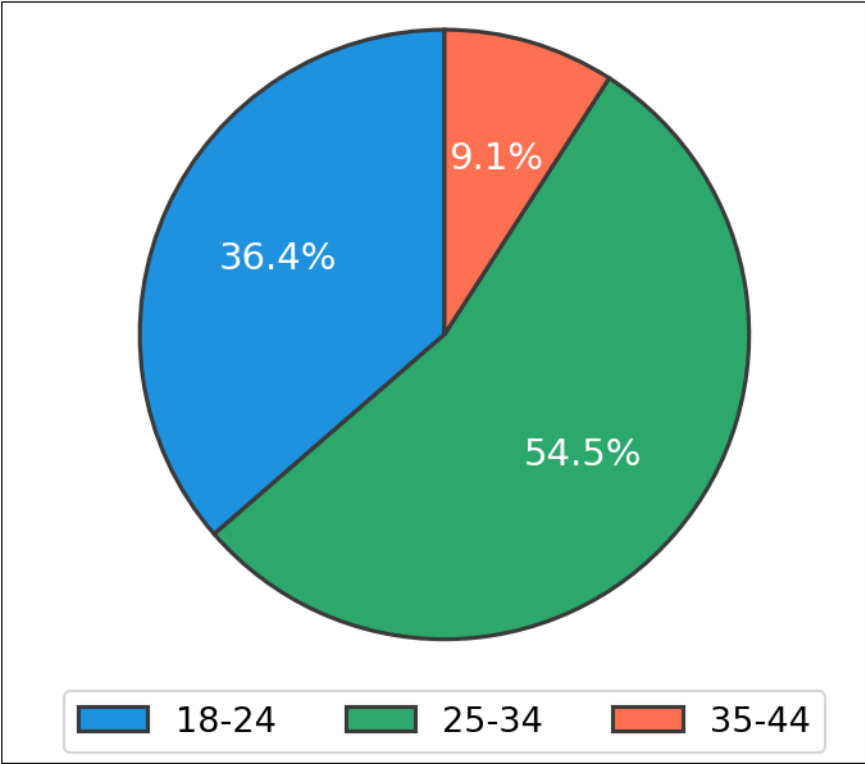
Use case: * Monitoring x v

Shelf: Select... v

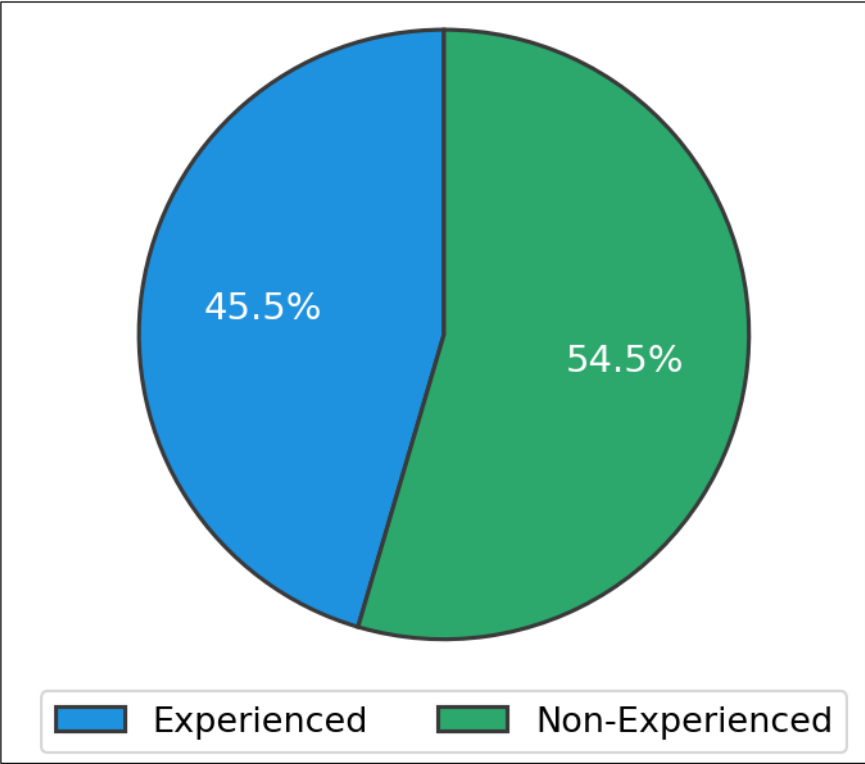
Location Detail: Select... v

Save sample Cancel

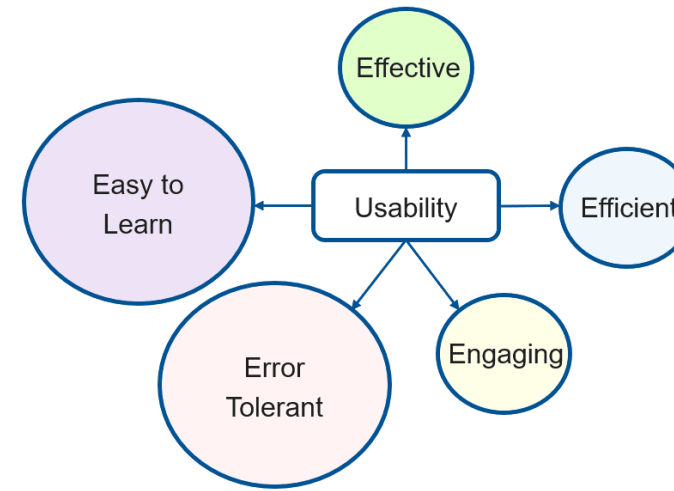
Age Group



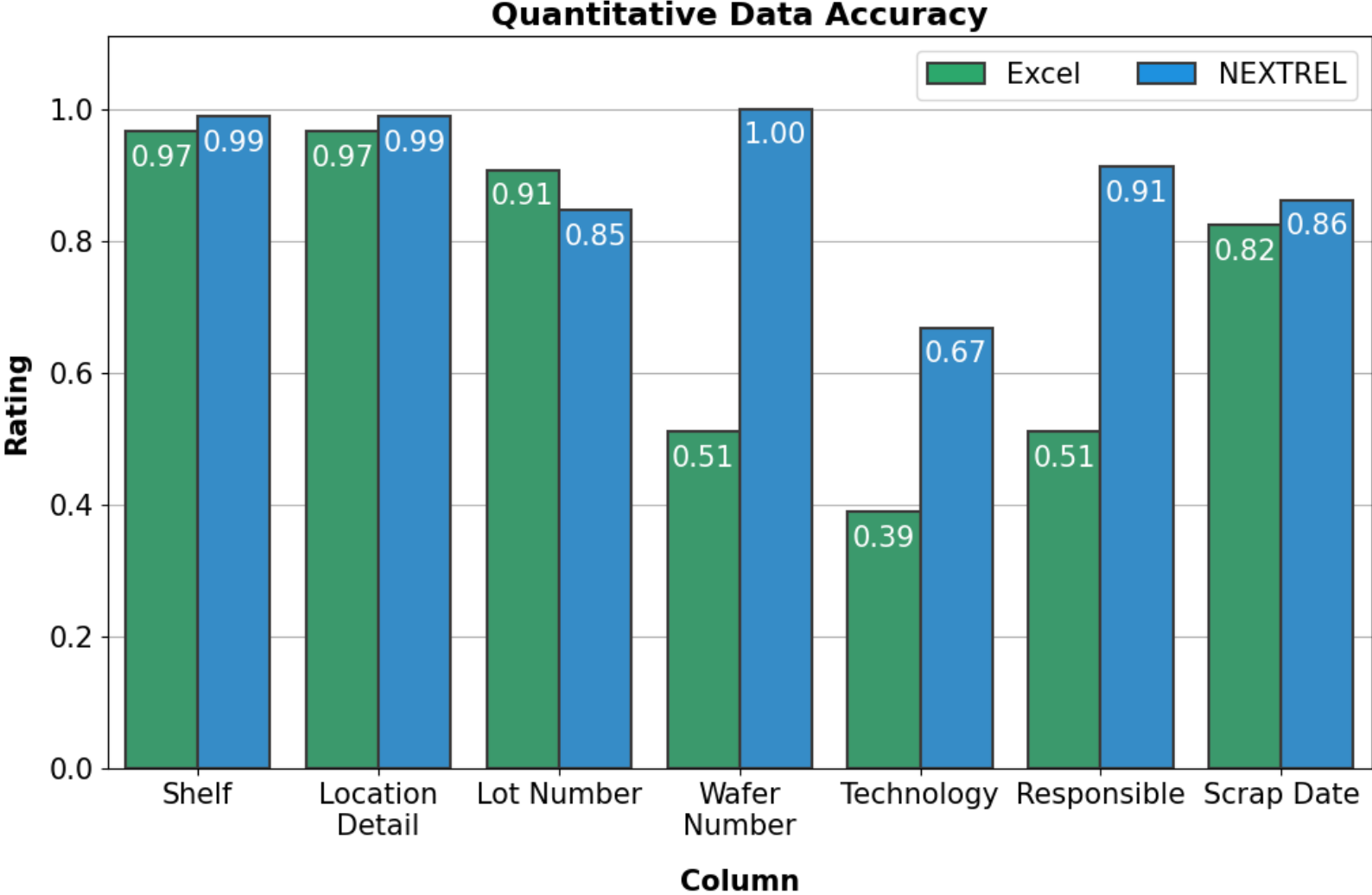
Experience Level

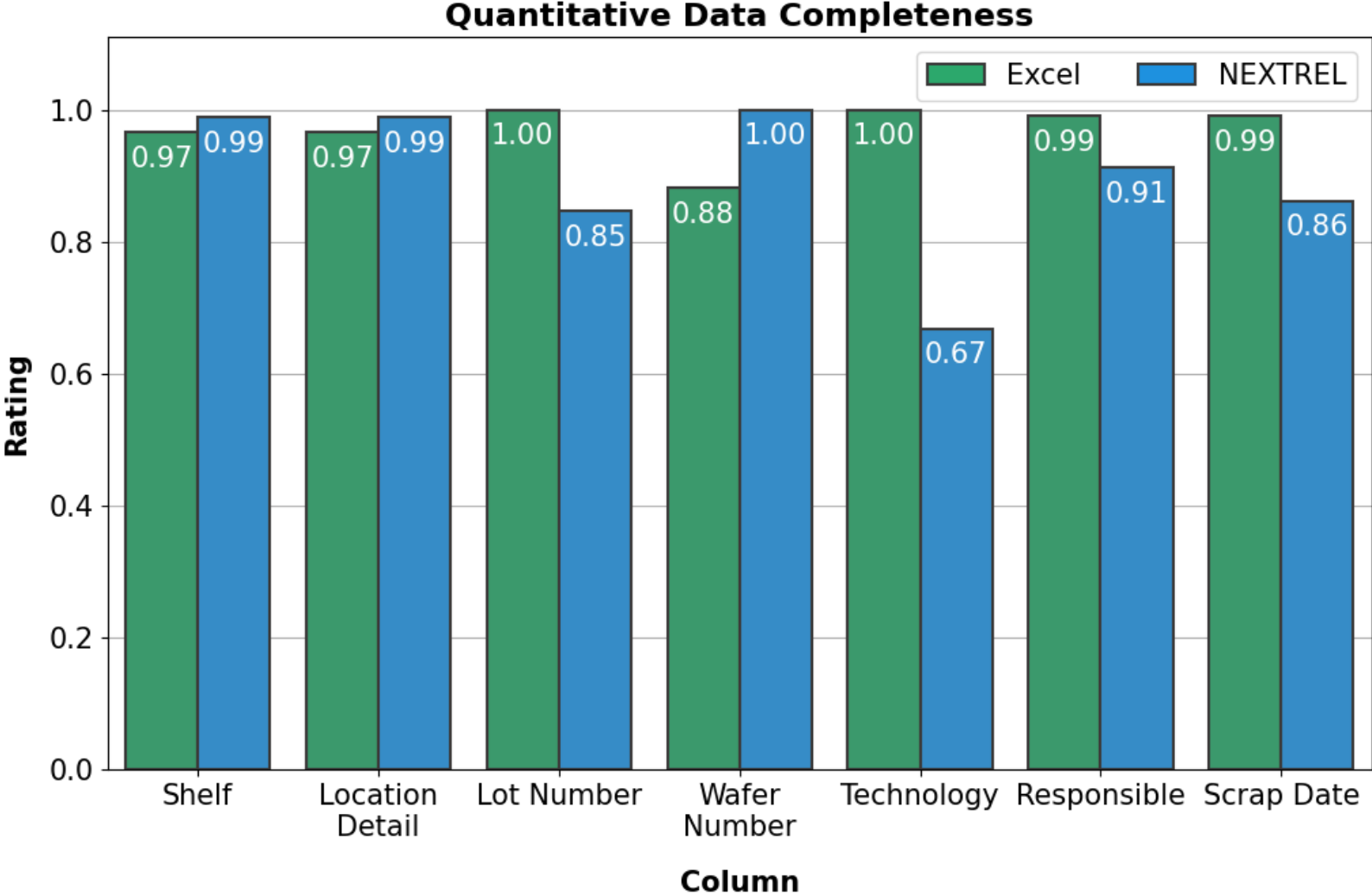


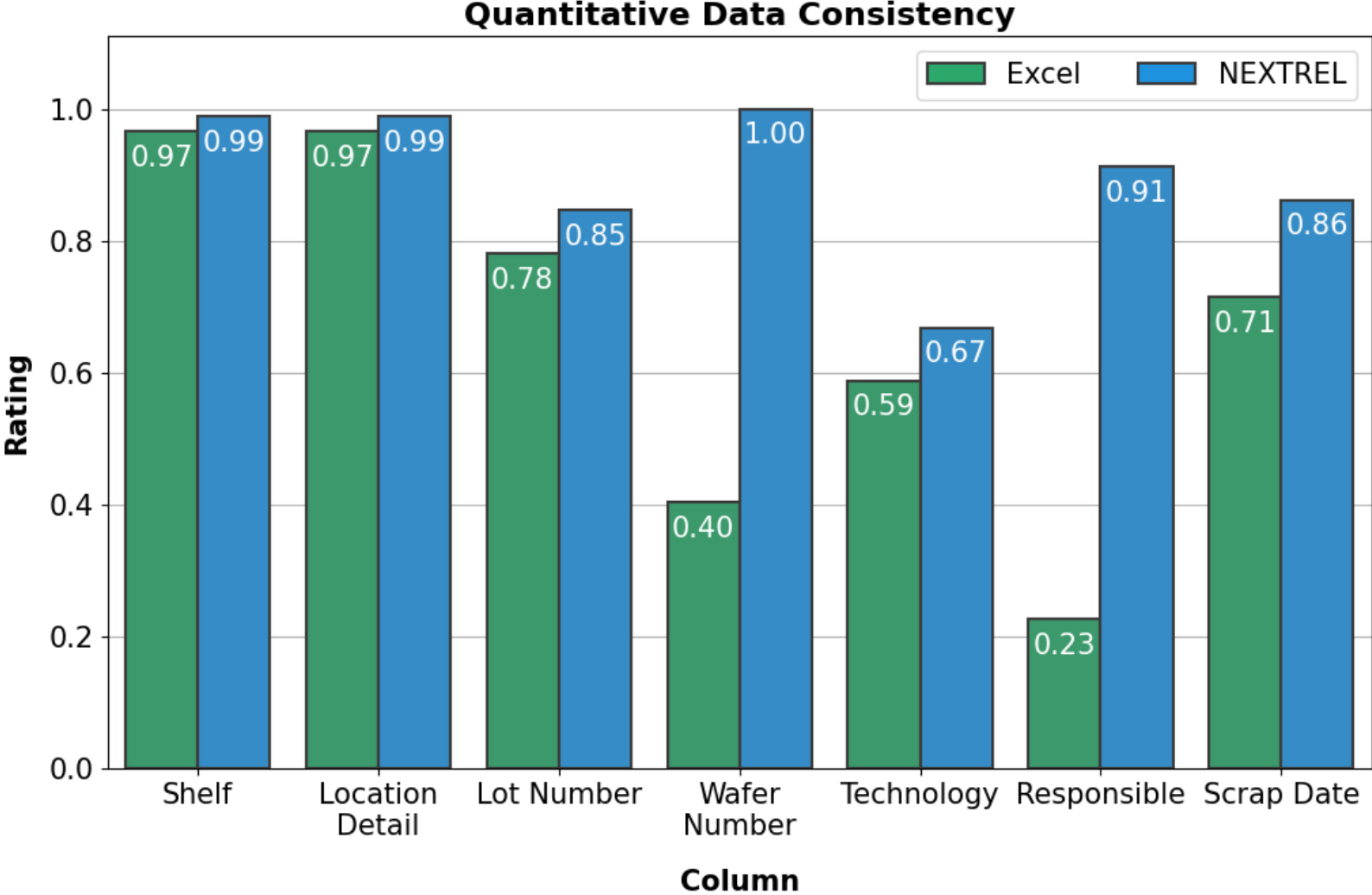
1. I was satisfied with this tool overall.	Strongly disagree						Strongly agree
		1	2	3	4	5	
2. I found this tool unnecessarily complex.	Strongly disagree						Strongly agree
		1	2	3	4	5	
3. I found it easy to fill in the data.	Strongly disagree						Strongly agree
		1	2	3	4	5	
4. I think that I would need the support of a technical person to be able to use this tool.	Strongly disagree						Strongly agree
		1	2	3	4	5	
5. I had a clear understanding of where each data needs to be entered.	Strongly disagree						Strongly agree
		1	2	3	4	5	
6. I found it easy to understand the correct format for entering the data.	Strongly disagree						Strongly agree
		1	2	3	4	5	
7. I received support from the tool for entering the data.	Strongly disagree						Strongly agree
		1	2	3	4	5	



Effective – ~14.3%
Efficient – ~14.3%
Engaging – ~14.3%
Error Tolerant – ~28.6%
Easy to Learn – ~28.6%







- Selection of data from *Excel* shown to participants

Shelf	Location Detail	Lot Number	Wafer Number	Technology	Responsible	Scrap Date
		TU889614.03	1,2,3	HSTF1200	Turan	Dec-24
w3		2 TU223344	7,8,3	CIT110305	Turan	Sep-25
w3		2 TU911312		CIT120310	Turan	Mar 27
w3		2 TU911313		CIT120310	Turan	Mar 27
w3		2 TU911314		CIT120310	Turan	Mar 27
w3		2 TU911315		CIT120310	Turan	Mar 27
w3		2 TU911316		CIT120310	Turan	Mar 27
w3		2 QU778812.04		C9LMG	Turan	Mar 27
w3		2 QU778812.05		C9LMG	Turan	Mar 27
w3		2 QU778812.06		C9LMG	Turan	Mar 27
w3		2 QU778812.07		C9LMG	Turan	Mar 27
w3		2 QU778812.08		C9LMG	Turan	Mar 27
w3		2 PL887026		P90TPA	Turan	Mar 27
w3		2 P99667 UWU		N	Turan	Mar 27
w3		2 GP2269622		CPTZ	Turan	Mar 27
w3		2 GP3696222		CPTZ	Turan	Mar 27
w3		2 GP2469622		CPTZ	Turan	Mar 27
w3		2 QU123444.01		C9LMG_GG	Turan	Mar 27
w3		2 QU123444.02		C9LMG_GG	Turan	Mar 27
w3		4 HF151005.00	4,5,6,7,8,9,10,11,12	MOSFET - 10 V	Turan	Aug-26
w3		2 HF888000.02		MOSFET - 15 V	Turan	Mar 27
w3		2 QU106700		PMOS99R (GaN)	Turan	Mar 27
w3		2 QU106701		PMOS99R (GaN)	Turan	Mar 27
w3		2 QU106702		PMOS99R (GaN)	Turan	Mar 27
w3		2 QU106703		PMOS99R (GaN)	Turan	Mar 27
w3		2 6ABB87HQZ2		CARD_7	Giebel, Turan	Mar 27
w3		2 6ABB87HQZ3		CARD_7	Giebel, Turan	Mar 27

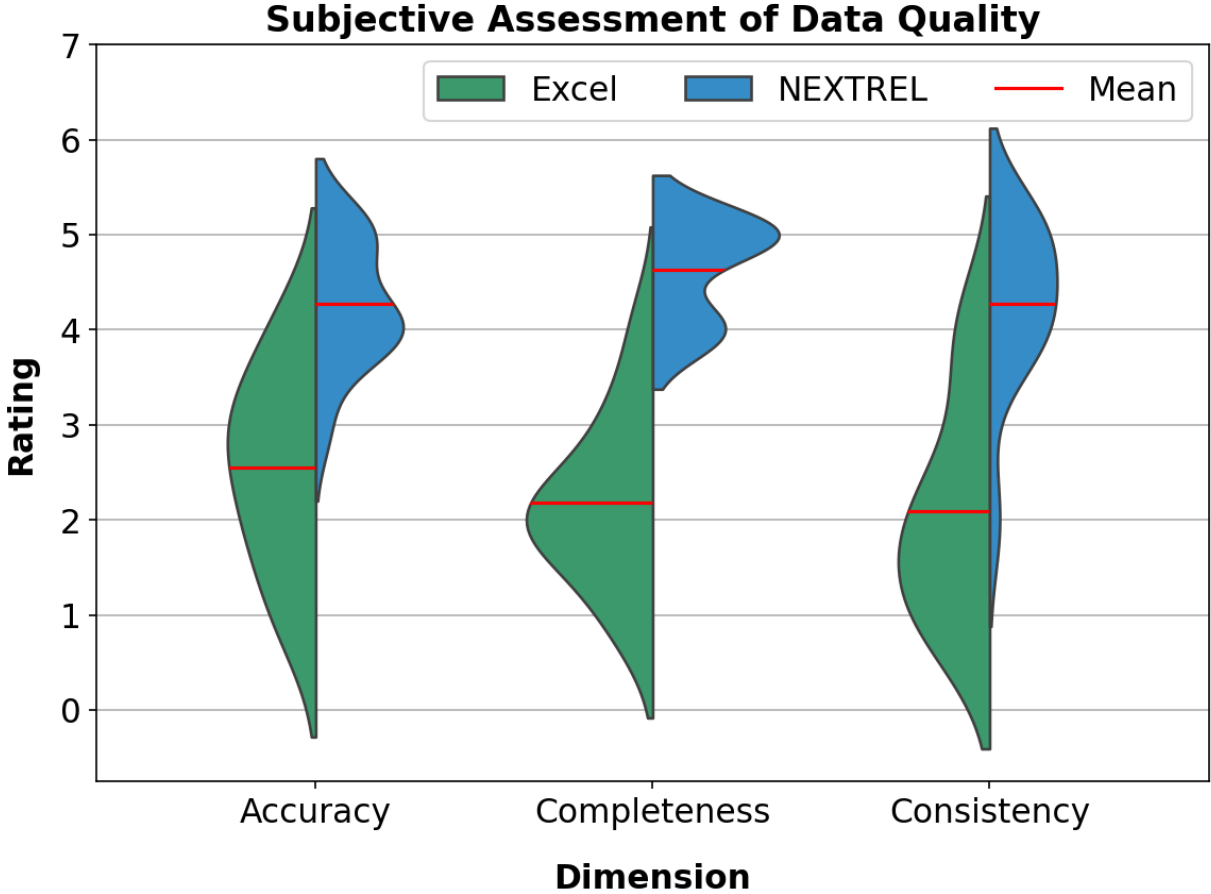
*data is altered for confidentiality purposes

- Selection of data from *NEXTREL* shown to participants

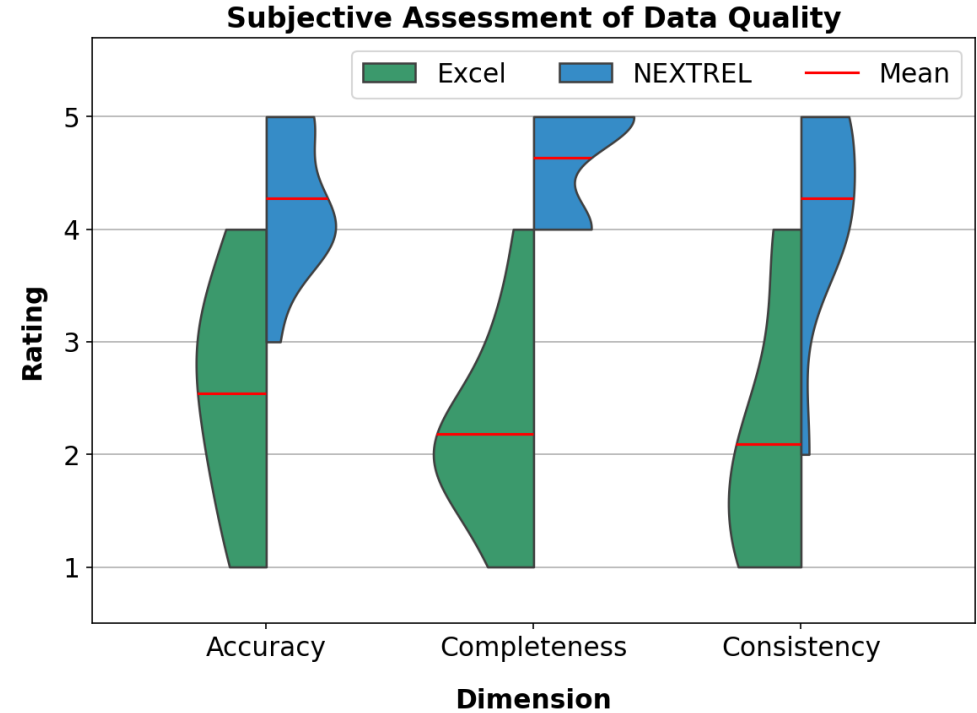
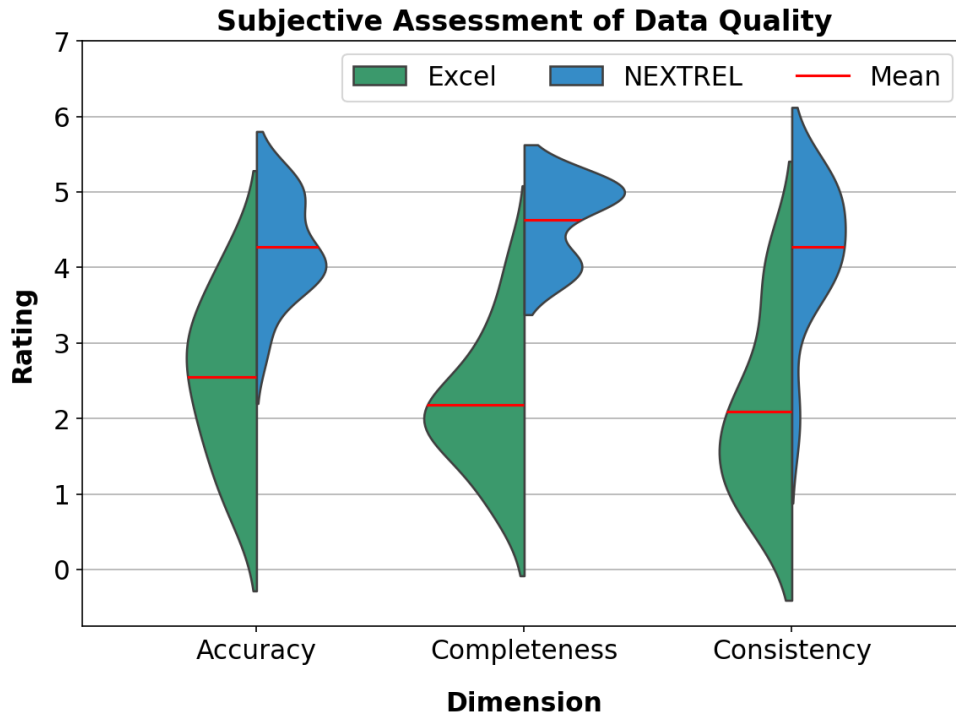
☰	Sample Type ▾	Lot Number ▾	Wafer Number ▾	Technology ▾	Status ▾
1	Wafer	TU180333	3	CIT001	In Storage
2	Wafer	TU180333	4	CIT001	In Storage
3	Wafer	TU180333	5	CIT001	In Storage
4	Wafer	TU180333	6	CIT001	In Storage
5	Wafer	TU180333	7	CIT001	In Storage
6	Wafer	TU180333	8	CIT001	In Storage
7	Wafer	TU180333	9	CIT001	In Storage
8	Wafer	TU285748	1	CIT002	In Storage
9	Wafer	TU285748	2	CIT002	In Storage
10	Wafer	TU285748	3	CIT002	In Storage
11	Wafer	TU285748	4	CIT002	In Storage
12	Wafer	TU285748	5	CIT002	In Storage
13	Wafer	TU285748	6	CIT002	In Storage
14	Wafer	TU285748	7	CIT002	In Storage
15	Wafer	TU285748	8	CIT002	In Storage
16	Wafer	TU285748	9	CIT002	In Storage
17	Wafer	TU285748	10	CIT002	In Storage

*data is altered for confidentiality purposes

- Violin plot with no data range restriction

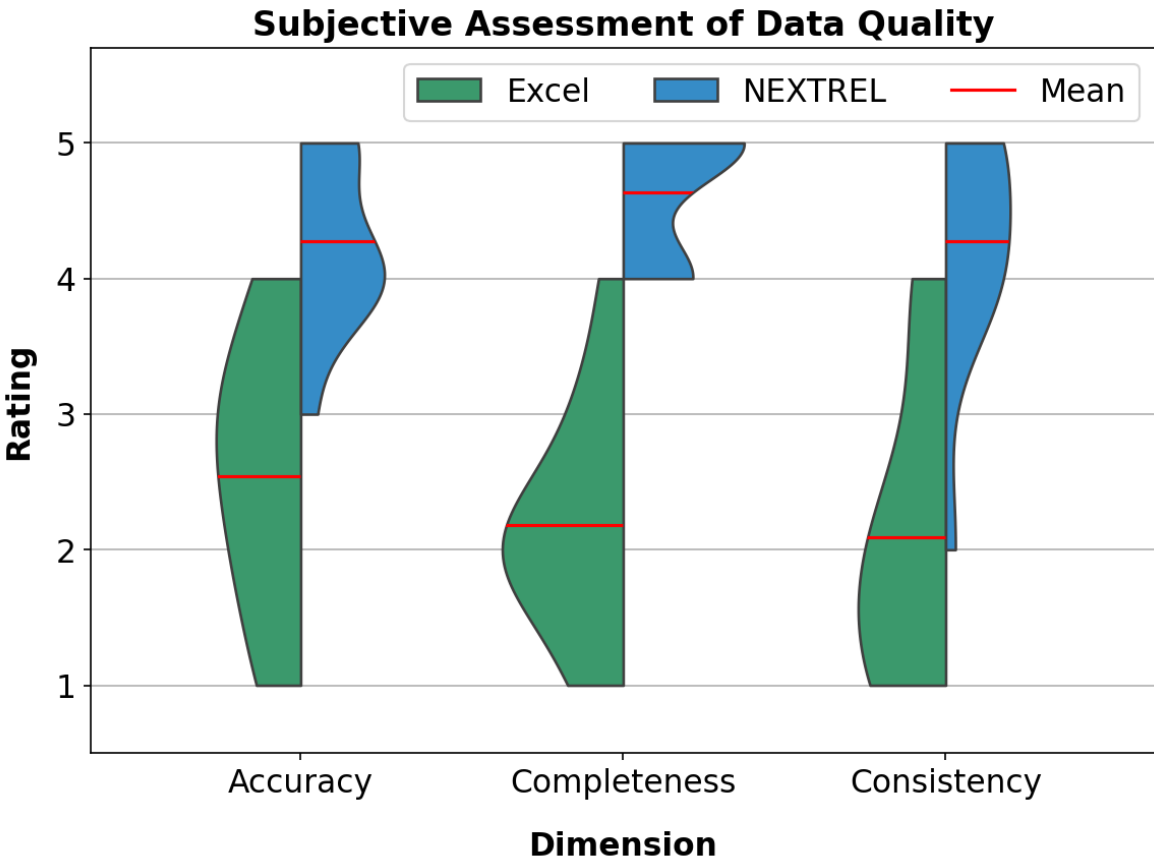
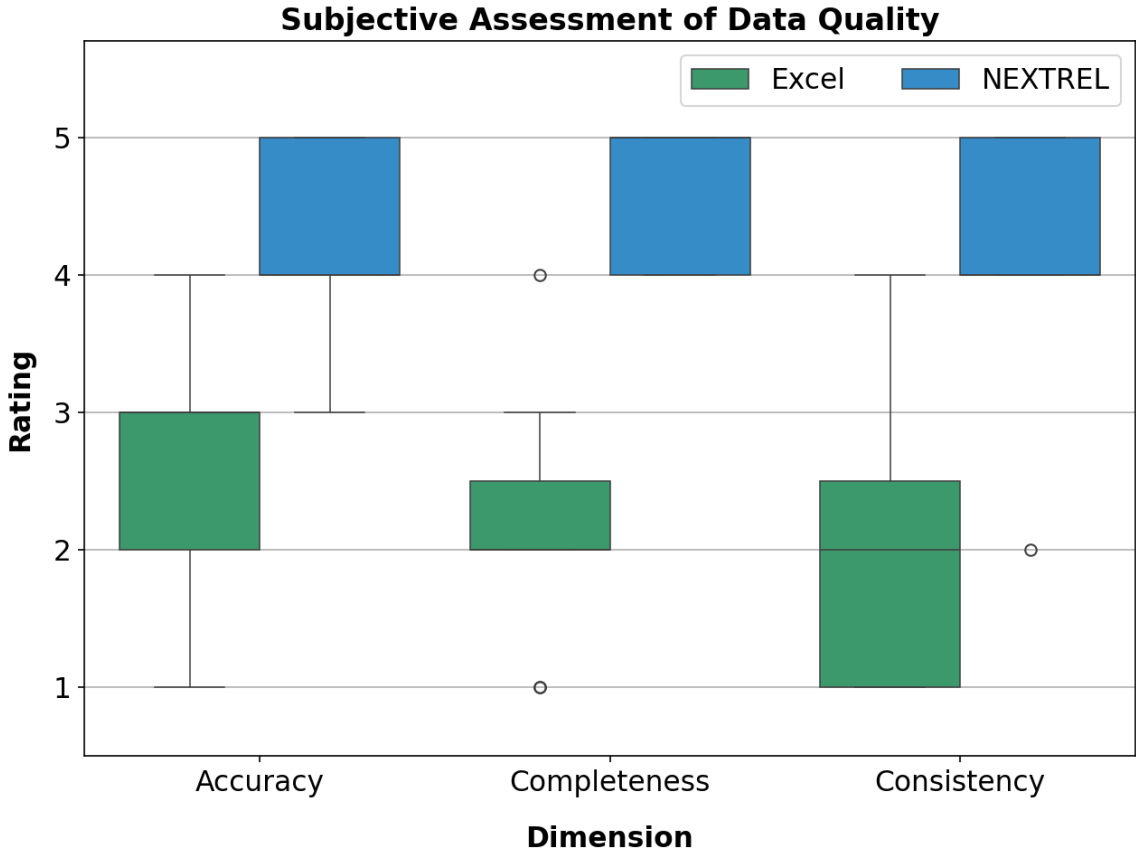


- Restricting violin plot to be within data range

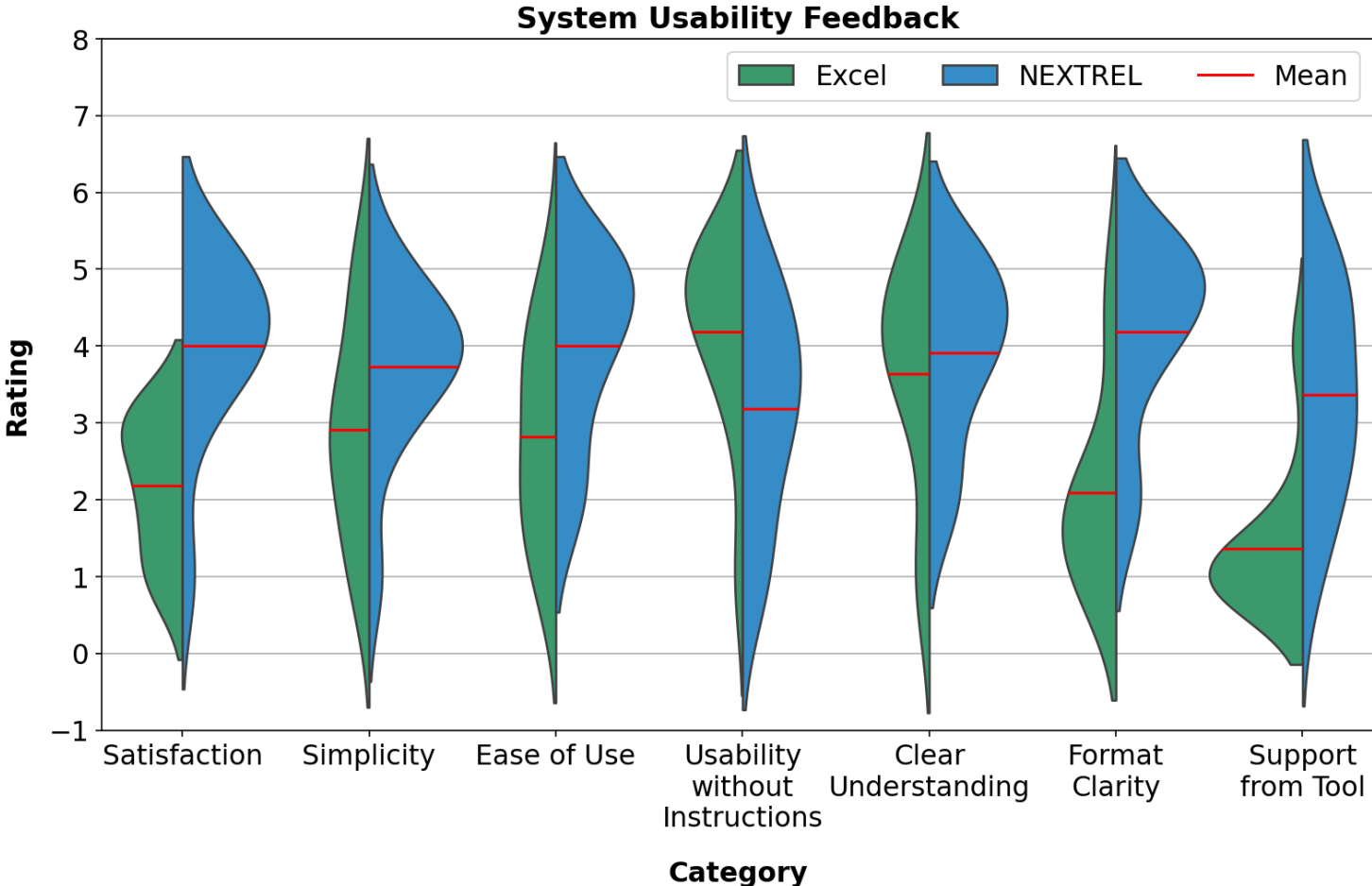


*both have different ranges of limit on y-axis

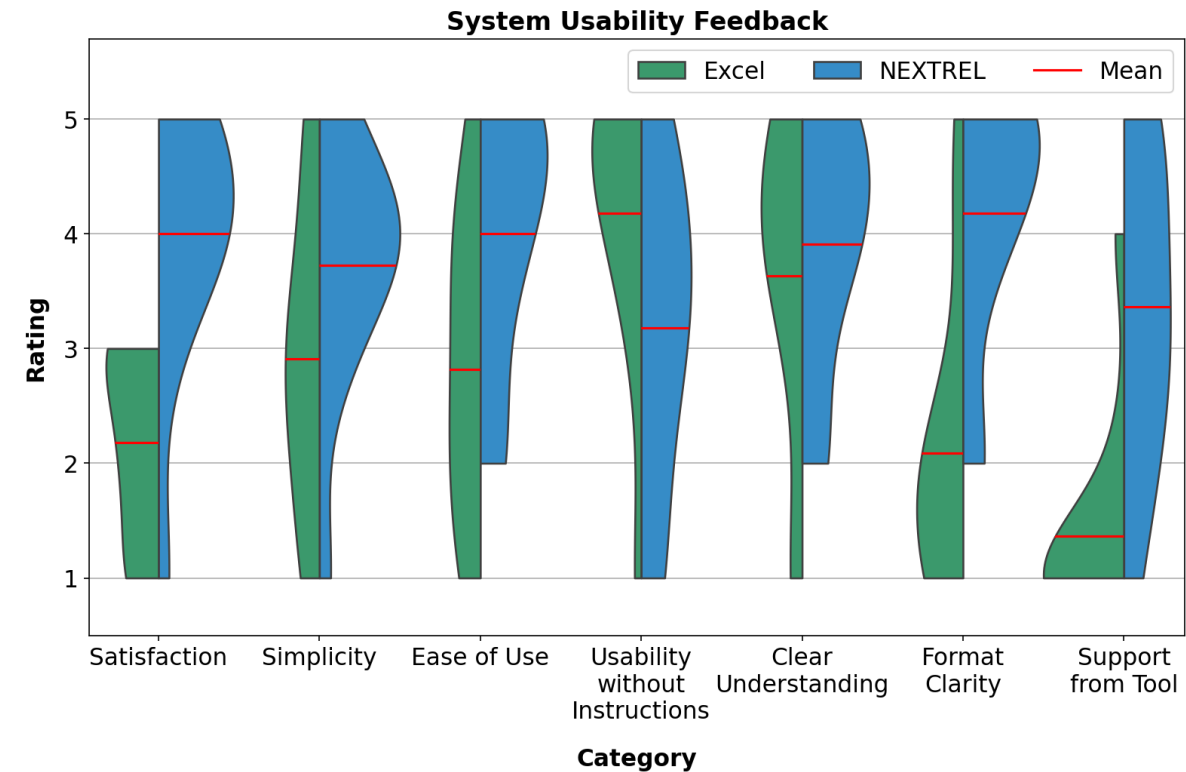
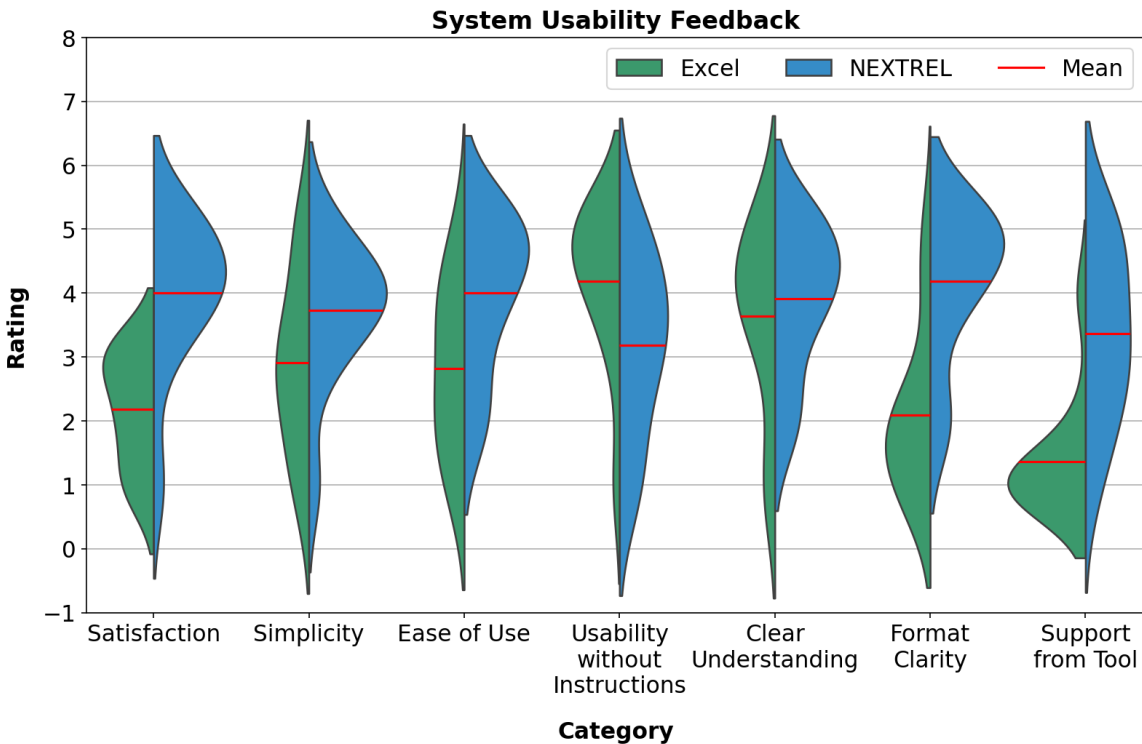
- Box plot versus violin plot



- Violin plot with no data range restriction



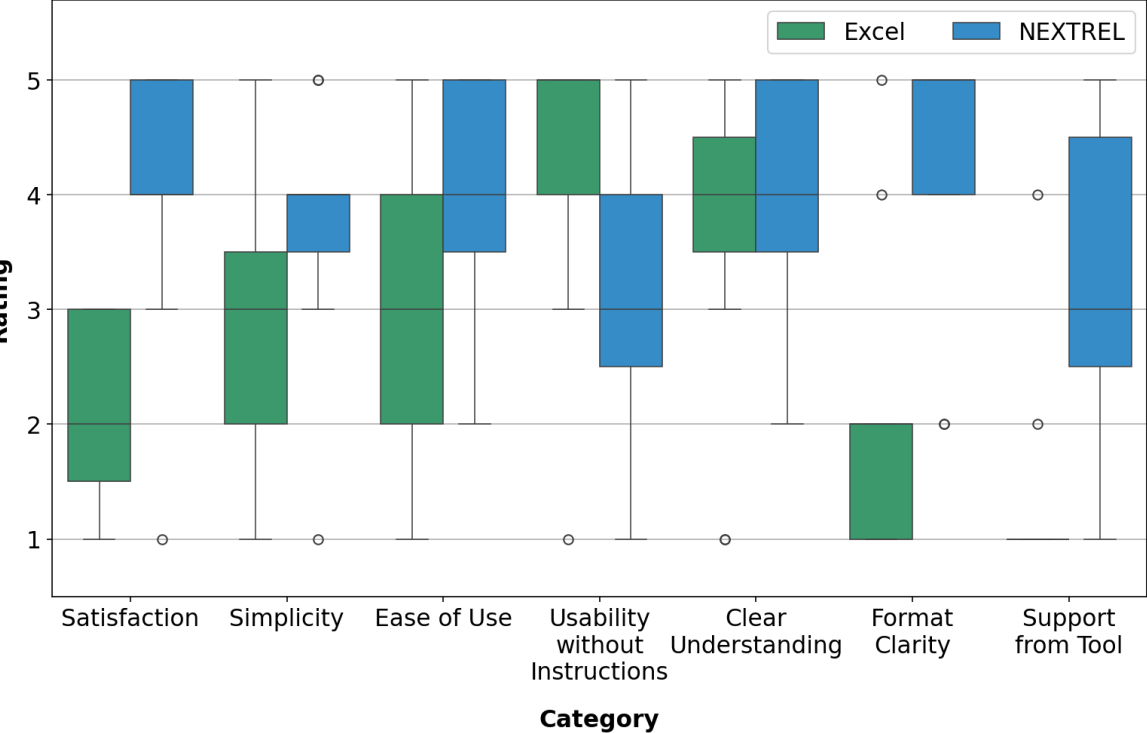
- Restricting violin plot to be within data range



*both have different ranges of limit on y-axis

- Box plot versus violin plot

System Usability Feedback



System Usability Feedback

