

Comparison ODK forms - demo

H. Langet

2023-09-26

Table of contents

SPA exit interview comparisons	2
Generic comparison	2
Comparison based on variable names	2
RCT Day 0 form comparison	4
Comparison based on variable names	4
Inter-form tests	9
Test for choices	12

List of Tables

1	SPA exit interview - Senegal vs. Tanzania - Added variable names	3
2	SPA exit interview - Senegal vs. Tanzania - Variable names with major label modifications	3
3	SPA exit interview - Senegal vs. Tanzania - Variable names with minor label modifications	4
4	Main Day 0 - Senegal vs. Tanzania - Added variable names	5
5	Main Day 0 - Senegal vs. Tanzania - Removed variable names	6
6	Main Day 0 - Senegal vs. Tanzania - Variable names with major label modifications	7
7	Main Day 0 - Senegal vs. Tanzania - Variable names with minor label modifications	8
8	Main Day 0 - Senegal vs. Tanzania - Variable names with type modifications	8

9	Tanzania - Main Day 0 vs. SPA exit interview - Added variable names	9
10	Tanzania - Main Day 0 vs. SPA exit interview - Removed variable names	11

```
root = r"C:\Users\langhe\switchdrive\Institution\TIMCI\1 Project details\11 DM\3 ODK\odk_forms"
```

SPA exit interview comparisons

```
in_xlsx1 = os.path.join(root, r"06 Tanzania\2-SPA-timeflow\02-TIMCI-SPA-CGEI-ref-form\02-TIMCI-SPA-CGEI-ref-form.xlsx")
in_xlsx2 = os.path.join(root, r"05 Senegal\2-SPA-timeflow\02-TIMCI-SPA-CGEI-ref-form\02-TIMCI-SPA-CGEI-ref-form.xlsx")
f1 = form.Form(in_xlsx1,
                survey = "SPA exit interview")
f2 = form.Form(in_xlsx2,
                survey = "SPA exit interview")
```

Generic comparison

```
print(f1.compare(f2))
```

Form ID is identical: 02-TIMCI-SPA-cgei
 Versions are different: 0.445 and 0.441
 Default languages are different: Swahili (sw) and Français (fr)

Comparison based on variable names

- Note:
- Identified added questions are displayed with the variable name and label of the closest label match in the reference XLSForm.
 - Identified removed questions are displayed with the variable name and label of the closest label match in the target XLSForm.

```
f1.detectAddedQuestions(f2)
```

Table 1: SPA exit interview - Senegal vs. Tanzania - Added variable names

	row	name	label	name_of_closest_lbl	closest_lbl	matching_score
0	49	i4_3	given paper record take completing referral	i4_4o		0.5
1	63	h4_2	show medicines prescriptions received	i4_4o		0.5
2	112	m3_2	how many children age 5 years currently live h...	i4_4o		0.5
3	116	m3_5	type stove use cooking household	i4_4o		0.5
4	117	m3_5o		i4_4o		1.0

```
f1.detectDeletedQuestions(f2)
```

```
major, minor = f1.detectModifiedLabels(f2)
```

```
major
```

Table 2: SPA exit interview - Senegal vs. Tanzania - Variable names with major label modifications

	name	row1	label1	row2	label2
0	i4_4o	36		52	
1	i4_6co	42		58	
2	j4_1d	56		73	
3	b1_8o	93		95	
4	b2_9c	98		100	
5	m3_3o	111		114	

```
minor
```

Table 3: SPA exit interview - Senegal vs. Tanzania - Variable names with minor label modifications

	name	row1	label1		row2	label2
0	m3_3	110	type toilet main toilet household members use	113	k4	type toilet main toilet household members use

```
f1.detectModifiedTypes(f2)
```

RCT Day 0 form comparison

```
in_xlsx3 = os.path.join(root, r"06 Tanzania\1-RCT-LS\01a-TIMCI-CRF-Facility-ref-form\01a-TIMCI-CRF-Facility-ref-form.xlsx")
in_xlsx4 = os.path.join(root, r"05 Senegal\1-RCT-LS\01a-TIMCI-CRF-Facility-Senegal-form\01a-TIMCI-CRF-Facility-Senegal-form.xlsx")
f3 = form.Form(in_xlsx3,
               survey = "RCT/LS day 0 form")
f4 = form.Form(in_xlsx4,
               survey = "RCT/LS day 0 form")

## Generic comparison

print(f3.compare(f4))
```

Form ID is identical: 01a-TIMCI-CRF-Facility

Versions are different: 0.356 and 0.347

Default languages are different: Swahili (sw) and Français (fr)

Comparison based on variable names

```
f3.detectAddedQuestions(f4)
```

Table 4: Main Day 0 - Senegal vs. Tanzania - Added variable names

	row	name	label	name_of_closest_lbl	closest_lbl
0	118	consent_img	take picture signed consent	booklet_pic	take picture child booklet make sure child
1	172	a4_c_6	area living	m2_2	living
2	173	a4_c_7		a3_c_1o	
3	174	location_name	location name	a3_c_1o	
4	175	location_type	location type	a3_c_1o	
5	176	location_rurb	rural urban	a3_c_1o	
6	205	cp_start		a3_c_1o	
7	206	cp_start_dec		a3_c_1o	
8	207	cp_start_disp		a3_c_1o	
9	218	injection	receive injection	vaccine_injection	receive immunisation injection today
10	219	injection_types	type injection	a3_c_1o	
11	220	injection_typeso		a3_c_1o	
12	245	paracetamol_route	indicate administration route paracetamol	gentamicin_route	indicate administration route gentamicin
13	246	salbutamol_route	indicate administration route salbutamol	gentamicin_route	indicate administration route gentamicin
14	267	b2_3	estimated amount money spent transport - overn...	a3_c_1o	
15	330	i2_1a	type referral	i2_1b	name referral facility
16	338	h2_2ayn	treatments recorded facility registry	child_hfid_yn	registration number allocated facility regis
17	340	amoxicillin_hf	clamoxyl bactox alphamox	a3_c_1o	
18	341	penig_hf	benzylpenicillin benpen	benzathine	benzathine benzylpenicillin (benzathine p
19	342	ceftriaxone_hf	ceftriaxone rocephin	ceftriaxone	ceftriaxone rocephin
20	343	ciprofloxacin_hf	ciprofloxacin ciproxin cifran	ciprofloxacin	ciprofloxacin cipon
21	344	ciprofloxacin_route_hf	indicate administration route ciprofloxacin	ciprofloxacin_route	indicate administration route ciprofloxacin
22	345	gentamicin_hf	gentamicin gentalline	gentamicin	gentamicin (genta)
23	346	gentamicin_route_hf	indicate administration route gentamicin	gentamicin_route	indicate administration route gentamicin
24	347	metronidazol_hf	metronidazol flagyl	metronidazol	flagyl metronidazole (metro)
25	348	metronidazol_route_hf	indicate administration route metronidazol	metronidazol_route	indicate administration route metronidazol
26	349	ampicillin_hf	ampicillin totapen	ampicillin	ampicillin ampiclox
27	350	azithromycin_hf	azithromycin zithromax	azithromycin	azithromycin zithromax
28	351	benzathine_hf	benzathine benzylpenicillin (benzathine penic...	benzathine	benzathine benzylpenicillin (benzathine p
29	352	aclav_hf	co-amoxiclav aclav soclav augmentin	aclav	amoxicillin + clavulanic acid amoxiclav co
30	353	cotrimoxazole_hf	co-trimoxazole septrin bactrim	a3_c_1o	
31	354	antibiotics_hf	indicate following antibiotics listed facility...	antibiotics	indicate following antibiotics among items
32	355	antimalarials_hf	indicate following antimalarials listed facili...	antimalarials	indicate following antimalarials among iter

	row	name	label	name_of_closest_lbl	closest_lbl
33	356	artesianate_route_hf	indicate administration route artesunate	artesianate_route	indicate administration route artesunate
34	357	quinine_route_hf	indicate administration route quinine	quinine_route	indicate administration route quinine
35	358	h2_2a	select drugs treatments listed facility registry	h2_2	select treatments among items listed drug
36	359	h2_2ao		a3_c_1o	
37	360	paracetamol_route_hf	indicate administration route paracetamol	gentamicin_route	indicate administration route gentamicin
38	361	salbutamol_route_hf	indicate administration route salbutamol	gentamicin_route	indicate administration route gentamicin
39	393	f2_5	haemoglobin level (g/dl)	f2_5dec	haemoglobin level (g/dl)

```
f3.detectDeletedQuestions(f4)
```

Table 5: Main Day 0 - Senegal vs. Tanzania - Removed variable names

	row	name	label	name_of_closest_lbl	closest_lbl	matching_score
0	13	district1	select current district	a3_c_1o		0.500000
1	14	fcode	select current facility	a3_c_1o		0.500000
2	128	arm	intervention	a3_c_1o		0.500000
3	137	contact_start		a3_c_1o		0.500000
4	138	contact_start_dec		a3_c_1o		0.500000
...
69	425	c3_3a	tb status recorded	c3_4	sickle-cell status recorded	0.543211
70	427	c3_6a	disabilities recorded	a3_c_1o		0.500000
71	431	c3_6	disability recorded	a3_c_1o		0.500000
72	432	c3_6o		a3_c_1o		1.000000
73	435	varia	thing would like mention regarding child/visit...	a3_c_1o		0.500000

```
major, minor = f3.detectModifiedLabels(f4)
```

```
major
```

Table 6: Main Day 0 - Senegal vs. Tanzania - Variable names with major label modifications

	name	row1	label1	row2	label2
0	a3_c_lo	81		75	
1	a4_a_11o	144		166	
2	a4_c_10	145	second phone number reached	167	alternative phone number reached
3	a4_c_9a	147	mention name village leader (focal point cont...	169	focal point within community contact
4	a4_c_9	148	mention name balozi (focal point contact with...	170	indicate name focal point contact within commu...
5	a4_c_4	149	division living	171	area / village living
6	c1_a_11o	217		201	
7	b2_2a_o	226		256	
8	b2_2b_o	231		259	
9	b1_4o	233		261	
10	b2_lo	236		264	
11	m2_lo	246		273	
12	m1_lo	250		277	
13	referral_note	291	given paper record take completing referral	216	referral note
14	h2_1a	296	healthcare provider give drugs treatment presc...	221	healthcare provider give drug prescription
15	h2_1b	297	take picture drugs/prescription	222	take picture drug prescription/prescribed medi...
16	h2_1c	298	take picture prescription make sure child name...	223	take picture drug prescription/prescribed medi...
17	b1_9	305	type visit according registry healthcare provider	211	type visit
18	amoxicillin	314	amoxil/amoxillin dt	225	ou clamoxyl ou bactox ou alfamox ou teramox ou...
19	aclav	315	amoxicillin + clavulanic acid amoxiclav co-amox...	237	ou co-amoxiclav ou amoxiclav ou soclav ou augm...
20	ampicillin	316	ampicillin ampiclox	234	ou totapen
21	benzathine	319	benzathine benzylpenicillin (benzathine penic...	236	benzathine-benzylpénicilline ou extencilline o...
22	penig	320	penicillin g/x-pen	226	benzylpenicilline ou pénig
23	ceftriaxone	321	ceftriaxone rocephin	227	ceftriaxone ou rocéphine
24	ciprofloxacin	322	ciprofloxacin cipon	228	ciprofloxacin ou ciproxin ou cifran
25	cotrimoxazole	324	septrin	238	ou bactrim
26	metronidazol	325	flagyl metronidazole (metro)	232	metronidazole ou flagyl
27	gentamicin	327	gentamicin (genta)	230	gentamicine ou gentalline
28	antibiotics	329	indicate following antibiotics among items lis...	239	indicate following antibiotics listed prescrip...
29	antimalarials	330	indicate following antimalarials among items l...	240	indicate following antimalarials listed prescr...
30	h2_2	336	select treatments among items listed drug pres...	243	select drugs/treatments listed prescription
31	h2_2o	337		244	
32	rxsevere	345	severe disease / severe febrile disease	305	severe disease / severe febrile disease / dang...

	name	row1	label1	row2	label2
33	g3_lo	364		327	
34	i2_1a1o	368		332	
35	i2_lo	370		334	
36	e2_2a	380	spo2 value	370	spo2 value (%)
37	f2_lo	401		389	
38	f2_10a	409	choose unit blood glucose value	395	blood glucose value (mg/dl)

minor

Table 7: Main Day 0 - Senegal vs. Tanzania - Variable names with minor label modifications

	name	row1	label1	row2	label2
0	a4_a_1	171	first name child	138	first names child
1	a4_a_5_1	184	first name child	150	first names child
2	azithromycin	318	azithromycin zithromax	235	azithromycine ou zithromax
3	metronidazol_route	326	indicate administration route metronidazole	233	indicate administration route metronidazol
4	vaccine_injection	338	receive immunisation injection today	247	receive immunisation injection

f3.detectModifiedTypes(f4)

Table 8: Main Day 0 - Senegal vs. Tanzania - Variable names with type modifications

	name	row1	type1	row2	type2
0	a3_a_11	74	select_one YESNO_FULLL	68	select_one YESNO_UNKN
1	a3_c_3a	86	select_one YESNO_DECL	80	select_one YESNO
2	ado_mother	89	select_one YESNO_UNKN	83	select_one YESNO
3	a4_a_10	142	text	164	integer
4	a4_c_10	145	text	167	integer
5	a4_c_4	149	select_one_from_file divisions.csv	171	select_one_from_file areas_villages.csv
6	a4_c_12	188	text	153	integer
7	referral_note	291	select_one YESNO	216	select_one YESNO_DECL

	name	row1	type1	row2	type2
8	antibiotics	329	select_multiple ANTIMICROBIALS	239	select_multiple ANTIBIOTICS
9	vaccine_injection	338	select_one YESNO	247	select_one YESNO_FULL
10	d2_6b	386	decimal	375	text
11	f2_10a	409	select_one Blood_Glucose_unit	395	integer

Inter-form tests

```
f3.detectAddedQuestions(f1)
```

Table 9: Tanzania - Main Day 0 vs. SPA exit interview - Added variable names

	row	name	label	name_of_closest_lbl	closest_lbl	match
0	15	fname	facility name	b1_1a	name facility referred	0.6190
1	17	ccode	country code	a3_c_1o		0.5000
2	26	i4_1	provider refer child	provider_name	name provider seen child	0.5388
3	27	h4_1	provider give prescribe medicines child take home	a3_c_1o		0.5000
4	28	g5_1	provider tell illness child	a3_c_1o		0.5000
5	33	i4_2	when need complete referral	a3_c_1o		0.5000
6	34	i4_5	told why go	a3_c_1o		0.5000
7	35	i4_4	told where go	a3_c_1o		0.5000
8	36	i4_4o		a3_c_1o		1.0000
9	37	i4_6	intend	a3_c_1o		0.5000
10	38	i4_6o	intend	a3_c_1o		0.5000
11	39	i4_6a	type facility go	a3_c_1o		0.5000
12	40	i4_6b	where look treatment	a3_c_1o		0.5000
13	41	i4_6c	ask advice	a3_c_1o		0.5000
14	42	i4_6co		a3_c_1o		1.0000
15	47	h4_3	provider explain how give medicines child home	a3_c_1o		0.5000
16	48	h4_4	how confident feel how much medication give da...	a3_c_1o		0.5000
17	53	j4_2	informed signs / symptoms require bring child ...	a3_c_1o		0.5000
18	54	j4_2a	signs symptoms	a3_c_1o		0.5000
19	55	j4_1	intend sick child get completely better become...	a3_c_1o		0.5000

	row	name	label	name_of_closest_lbl	closest_lbl	match
20	56	j4_1d		a3_c_1o		1.0000
21	57	j4_1a	type facility go	a3_c_1o		0.5000
22	58	j4_1b	where look treatment	a3_c_1o		0.5000
23	59	j4_1c	ask advice	a3_c_1o		0.5000
24	60	h4_6	given general information advice feeding breas...	m2_9	how many months breastfeed additional feeding ...	0.5241
25	66	e4_2	explain why device used	a3_c_1o		0.5000
26	67	e4_3	provider explain result given device	a3_c_1o		0.5000
27	72	k2_1	provider use tablet like one consultation child	a3_c_1o		0.5000
28	73	k2_2	provider explain why using tablet	a3_c_1o		0.5000
29	74	k2_2a	explanation given also child	a3_c_1o		0.5000
30	75	k2_3	how feel fact provider used tablet consultatio...	a3_c_1o		0.5000
31	80	l3_1	how feel overall service received facility today	a3_c_1o		0.5000
32	81	l3_2	feel provider treated child respect	a3_c_1o		0.5000
33	82	l3_3	find provider kind	a3_c_1o		0.5000
34	83	l3_4	find provider showed concern empathy	a3_c_1o		0.5000
35	84	l3_5	provider speak language understand	a3_c_1o		0.5000
36	85	l3_6	service delayed kept waiting long time	a3_c_1o		0.5000
37	86	l3_7	would recommend facility friend / family sick ...	a3_c_1o		0.5000
38	91	b1_7	facility closest health facility home	a3_c_1o		0.5000
39	92	b1_8	main reason choose coming today rather going c...	a3_a_10	main reason bring child facility today	0.5373
40	93	b1_8o		a3_c_1o		1.0000
41	94	b2_10	miss work bring child facility today	a3_a_10	main reason bring child facility today	0.5623
42	95	b2_10a	how many work days miss result visit	a3_c_1o		0.5000
43	96	b2_9a	pay something facility today	a3_c_1o		0.5000
44	97	b2_9b	pay	a3_c_1o		0.5000
45	98	b2_9c		a3_c_1o		1.0000
46	99	b2_4a	estimated amount paid registration	a3_c_1o		0.5000
47	100	b2_5a	estimated amount paid consultation	a3_c_1o		0.5000
48	101	b2_6a	estimated amount paid laboratory investigations	a3_c_1o		0.5000
49	102	b2_7	estimated amount money spent treatment child (...)	a3_c_1o		0.5000
50	103	b2_8	intend buy medicines outside facility	a3_c_1o		0.5000
51	109	m3_1b	head household	a3_c_1o		0.5000

f3.detectDeletedQuestions(f1)

Table 10: Tanzania - Main Day 0 vs. SPA exit interview - Removed variable names

	row	name	label	name_of_closest_lbl	closest_lbl	matching_score
0	5	start_geopoint		i4_4o		0.5
1	6	screening_id	screening id	i4_4o		0.5
2	13	district1	select current district	b1_4	select current district	1.0
3	20	dobk	know date birth child	i4_4o		0.5
4	21	screening_start		i4_4o		0.5
...
302	429	report_end_dec		i4_4o		0.5
303	430	report_end_disp		i4_4o		0.5
304	431	c3_6	disability recorded	i4_4o		0.5
305	432	c3_6o		i4_4o		1.0
306	435	varia	thing would like mention regarding child/visit...	i4_4o		0.5

f3.detectSimilarLabels(f1)

	row1	name1	label1	type1	row2	name2	label2
0	13	district1	select current district	select_one_from_file districts.csv	11	b1_4	select
1	14	fcode	select current facility	select_one_from_file facilities.csv	12	fcode	select
2	81	a3_c_1o		text	36	i4_4o	
3	97	former_timci_id1	scan participants qr code	barcode	20	a1_a_4	scan p
4	98	former_timci_id2	qr code scanning possible , manually enter par...	text	21	a1_a_4a	qr coc
5	122	a1_a_4a	qr code scanning possible , manually enter par...	text	21	a1_a_4a	qr coc
6	128	arm	intervention	calculate	16	arm	interv
7	144	a4_a_11o		text	36	i4_4o	
8	150	a4_c_4_oth		text	36	i4_4o	
9	155	a4_c_4a_oth		text	36	i4_4o	
10	158	a4_c_4b_oth		text	36	i4_4o	
11	161	a4_c_4c_oth		text	36	i4_4o	
12	217	c1_a_11o		text	36	i4_4o	

	row1	name1	label1	type1	row2	name2	label2
13	226	b2_2a_o		text	36	i4_4o	
15	231	b2_2b_o		text	36	i4_4o	
16	233	b1_4o		text	36	i4_4o	
17	236	b2_1o		text	36	i4_4o	
18	246	m2_1o		text	36	i4_4o	
19	250	m1_1o		text	36	i4_4o	
20	269	m3_6	where households main source drinking water lo...	select_one WATER_SRC_LOC	113	m3_6	where
21	270	m3_7	how much time take household get main source d...	select_one TIME_2_SRC_LOC	114	m3_7	how m
22	271	m3_3	type toilet main toilet household members use	select_one MAIN_TOILET	110	m3_3	type t
23	272	m3_3o		text	36	i4_4o	
24	273	m3_4	toilet shared another household	select_one YESNO_DECL	112	m3_4	toilet
25	274	m3_8a	type floor home	select_one FLOOR_TYPES	115	m3_8a	type f
26	275	m3_9a	type roof home	select_one ROOF_TYPES	116	m3_9a	type r
27	280	m3_5o		text	36	i4_4o	
28	290	i2_1a1_cgo		text	36	i4_4o	
29	299	e4_1	provider use device represented following pict...	select_one YESNO_FULL	65	e4_1	provid
30	317	ampicillin_spec		select_one AMPICILLIN_DETAILS	36	i4_4o	
31	337	h2_2o		text	36	i4_4o	
32	364	g3_1o		text	36	i4_4o	
33	368	i2_1a1o		text	36	i4_4o	
35	370	i2_1o		text	36	i4_4o	
36	401	f2_1o		text	36	i4_4o	
37	432	c3_6o		text	36	i4_4o	
14	228	b1_1a	name facility referred	text	15	fname	facilit
34	369	i2_1b	name referral facility	select_one_from_file referral_facilities.csv	15	fname	facilit

Test for choices

```
#choices_df["list_name"].dropna().unique()
```