Analytic Guidelines:
Creating Disability Identifiers Using the Washington Group Short Set (WG-SS) SPSS Syntax

Introduction

Disability is best understood as a continuum. In terms of difficulty functioning, the 'difficulty' can be operationalized through a range of descriptors from no difficulty at all, through some difficulty and a lot of difficulty to completely unable to carry out the action. Each of these descriptors represents a cut-off or threshold in the determination of a final disability identifier; for example, to define those with and without disability. These levels of functioning are also represented in the response categories to the WG Short Set on Functioning (WG-SS).

Disability prevalence is not a single statistic, but can be calculated at various thresholds depending on the purposes of both data collection and reporting. For example, if the purpose is to provide for equitable access to public spaces – then the level of inclusion for a disability identifier might be *some difficulty*, since those with even minor levels of difficulty functioning would likely benefit from adaptations made to remove barriers and ease access. The installation of escalators in place of stairs, for instance, is a common universal design element that benefits persons with a wide range of mobility difficulty. Alternatively, if the purpose is to provide subsidies or allowances – the level of inclusion for a disability identifier might be *cannot do at all* since only those with more severe functional limitations would meet stricter eligibility criteria.

> This fifth in a series of **Washington Group Implementation Documents** covers analytic guidance, including SPSS syntax, when using the Washington Group Short Set questions. Guidance on creating disability identifiers is provided.
>
> Additional Implementation Documents cover: The Washington Group Tools; The Washington Group Short Set on Functioning; Translation of the WG Tools; Question Specifications; and other tools developed by the WG and partners for the measurement of disability.
>
> For more information, visit the Washington Group website: http://www.washingtongroup-disability.com/.

The SPSS syntax described in this document provides for the calculation of four disability identifiers at four thresholds. The population of those *with disability* using these four different thresholds produces the following four disability identifiers:

- **DISABILITY1**: the level of inclusion is at least one domain/question is coded SOME DIFFICULTY or A LOT OF DIFFICULTY or CANNOT DO AT ALL.

- **DISABILITY2**: the level of inclusion is at least 2 domains/questions are coded SOME DIFFICULTY or any 1 domain/question is coded A LOT OF DIFFICULTY or CANNOT DO AT ALL.

- **DISABILITY3**: the level of inclusion is any 1 domain/question is coded A LOT OF DIFFICULTY or CANNOT DO AT ALL.

NOTE: DISABILITY3 IS THE CUT-OFF RECOMMENDED BY THE WG.

- **DISABILITY4**: the level of inclusion is any one domain is coded CANNOT DO AT ALL (4).

NOTE: The SPSS syntax is based on the *variable labels and value labels* indicated in the tables below. Ensure that you use the same *variable and value labels* OR revise the SPSS syntax to reflect the *labels* used in your database.

The WG-SS Questions as they appear in the U.S. National Health Interview Survey (NHIS):

WG Short Set Questions/Domains	Variable Label
1. Do you have difficulty seeing even if wearing glasses?	VIS_SS
2. Do you have difficulty hearing even if using a hearing aid?	HEAR_SS
3. Do you have difficulty walking or climbing stairs?	MOB_SS
4. Do you have difficulty remembering or concentrating?	COG_SS
5. Do you have difficulty with (self-care such as) washing all over or dressing?	UB_SS
6. Using your usual language, do you have difficulty communicating (for example understanding or being understood by others)?	COM_SS

The value labels used for each of the WG-SS questions are:

1. No difficulty
2. Yes, some difficulty
3. Yes, a lot of difficulty
4. Cannot do at all
7. Refused
8. Not ascertained
9. Don't know

SPSS WG Short Set Syntax Annotated with Output Tables

Actual SPSS syntax is indented – commands in **bold text**.

NOTE: For data analysis, use your standard weighting and estimation techniques.

Step 1: Generate frequency distributions on each of the six domain variables.

The syntax below produces frequency distributions on each the six domains. Codes 7 (REFUSED), 8 (NOT ASCERTAINED) and 9 (DON'T KNOW) are INCLUDED.

http://www.washingtongroup-disability.com/

FREQUENCIES VIS_SS HEAR_SS MOB_SS COM_SS UB_SS COG_SS.

VIS_SS

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	No difficulty	13690	79.0	79.0	79.0
	Some difficulty	2708	15.6	15.6	94.6
	A lot of difficulty	333	1.9	1.9	96.6
	Cannot do at all	36	.2	.2	96.8
	Refused	21	.1	.1	96.9
	Not ascertained	532	3.1	3.1	100.0
	Don't know	6	.0	.0	100.0
	Total	17326	100.0	100.0	

HEAR_SS

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	No difficulty	13680	79.0	79.0	79.0
	Some difficulty	2753	15.9	15.9	94.8
	A lot of difficulty	310	1.8	1.8	96.6
	Cannot do at all	23	.1	.1	96.8
	Refused	24	.1	.1	96.9
	Not ascertained	534	3.1	3.1	100.0
	Don't know	2	.0	.0	100.0
	Total	17326	100.0	100.0	

MOB_SS

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	No difficulty	13424	77.5	77.5	77.5
	Some difficulty	2165	12.5	12.5	90.0
	A lot of difficulty	792	4.6	4.6	94.5
	Cannot do at all	380	2.2	2.2	96.7
	Refused	25	.1	.1	96.9
	Not ascertained	536	3.1	3.1	100.0
	Don't know	4	.0	.0	100.0
	Total	17326	100.0	100.0	

COM_SS

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	No difficulty	15874	91.6	91.6	91.6
	Some difficulty	745	4.3	4.3	95.9
	A lot of difficulty	94	.5	.5	96.5
	Cannot do at all	43	.2	.2	96.7
	Refused	25	.1	.1	96.9
	Not ascertained	543	3.1	3.1	100.0
	Don't know	2	.0	.0	100.0
	Total	17326	100.0	100.0	

UB_SS

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	No difficulty	16029	92.5	92.5	92.5
	Some difficulty	544	3.1	3.1	95.7
	A lot of difficulty	114	.7	.7	96.3
	Cannot do at all	68	.4	.4	96.7
	Refused	25	.1	.1	96.8
	Not ascertained	544	3.1	3.1	100.0
	Don't know	2	.0	.0	100.0
	Total	17326	100.0	100.0	

COG_SS

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	No difficulty	13719	79.2	79.2	79.2
	Some difficulty	2632	15.2	15.2	94.4
	A lot of difficulty	382	2.2	2.2	96.6
	Cannot do at all	20	.1	.1	96.7
	Refused	25	.1	.1	96.8
	Not ascertained	543	3.1	3.1	100.0
	Don't know	5	.0	.0	100.0
	Total	17326	100.0	100.0	

Step 2: Recode 7 (REFUSED), 8 (NOT ASCERTAINED) and 9 (DON'T KNOW) as MISSING and generate frequency distributions on each of the six domain variables.

The syntax below will yield domain-specific frequencies and thereby, prevalence of disability by domain of functioning. Codes 7, 8 and 9 are TEMPORARILY coded as MISSING.

Combine A LOT OF DIFFICULTY (3) or CANNOT DO AT ALL (4) for domain-specific prevalence at the WG recommended cut-off. (Highlighted in the tables.)

TEMPORARY.
RECODE VIS_SS HEAR_SS MOB_SS COM_SS UB_SS COG_SS (7 thru 9=SYSMIS).
FREQUENCIES VIS_SS HEAR_SS MOB_SS COM_SS UB_SS COG_SS.

VIS_SS

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	No difficulty	13690	79.0	81.6	81.6
	Some difficulty	2708	15.6	16.2	97.8
	A lot of difficulty	**333**	**1.9**	**2.0**	**99.8**
	Cannot do at all	**36**	**.2**	**.2**	**100.0**
	Total	16767	96.8	100.0	
Missing		559	3.2		
Total		17326	100.0		

HEAR_SS

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	No difficulty	13680	79.0	81.6	81.6
	Some difficulty	2753	15.9	16.4	98.0
	A lot of difficulty	**310**	**1.8**	**1.8**	**99.9**
	Cannot do at all	**23**	**.1**	**.1**	**100.0**
	Total	16766	96.8	100.0	
Missing		560	3.2		
Total		17326	100.0		

MOB_SS

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	No difficulty	13424	77.5	80.1	80.1
	Some difficulty	2165	12.5	12.9	93.0
	A lot of difficulty	**792**	**4.6**	**4.7**	**97.7**
	Cannot do at all	**380**	**2.2**	**2.3**	**100.0**
	Total	16761	96.7	100.0	
Missing		565	3.3		
Total		17326	100.0		

COM_SS

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	No difficulty	15874	91.6	94.7	94.7
	Some difficulty	745	4.3	4.4	99.2
	A lot of difficulty	**94**	**.5**	**.6**	**99.7**
	Cannot do at all	**43**	**.2**	**.3**	**100.0**
	Total	16756	96.7	100.0	
Missing		570	3.3		
Total		17326	100.0		

UB_SS

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	No difficulty	16029	92.5	95.7	95.7
	Some difficulty	544	3.1	3.2	98.9
	A lot of difficulty	**114**	**.7**	**.7**	**99.6**
	Cannot do at all	**68**	**.4**	**.4**	**100.0**
	Total	16755	96.7	100.0	
Missing		571	3.3		
Total		17326	100.0		

COG_SS

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	No difficulty	13719	79.2	81.9	81.9
	Some difficulty	2632	15.2	15.7	97.6
	A lot of difficulty	**382**	**2.2**	**2.3**	**99.9**
	Cannot do at all	**20**	**.1**	**.1**	**100.0**
	Total	16753	96.7	100.0	
Missing		573	3.3		
Total		17326	100.0		

Step 3: Calculate a variable, SUM_234

SUM_234 summates the number of domains coded SOME DIFFICULTY (2) or A LOT OF DIFFICULTY (3) or CANNOT DO AT ALL (4) for each person. This new variable is used in the determination of disability identifiers: DISABILITY1 and DISABILITY2.

The syntax below counts the number of domains/questions a person has that are coded SOME DIFFICULTY (2) or A LOT OF DIFFICULTY (3) or CANNOT DO AT ALL (4).

Possible range 0: no difficulties in any domain, to 6: all six domains coded SOME DIFFICULTY (2) or A LOT OF DIFFICULTY (3) or CANNOT DO AT ALL (4).

> **COUNT** SUM_234 = VIS_SS HEAR_SS MOB_SS COM_SS COG_SS UB_SS (2 thru 4).
> **FREQUENCIES** SUM_234.

SUM_234

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	.00	9815	56.6	56.6	56.6
	1.00	3839	22.2	22.2	78.8
	2.00	1892	10.9	10.9	89.7
	3.00	989	5.7	5.7	95.4
	4.00	481	2.8	2.8	98.2
	5.00	232	1.3	1.3	99.5
	6.00	78	.5	.5	100.0
	Total	17326	100.0	100.0	

Step 4: Calculate a variable, SUM_34

SUM_34 summates the number of domains coded A LOT OF DIFFICULTY (3) or CANNOT DO AT ALL (4) for each person. This new variable is used in the determination of disability identifier: DISABILITY2.

The syntax below counts the number of domains/questions a person has that are coded A LOT OF DIFFICULTY (3) or CANNOT DO AT ALL (4)

Possible range 0: no difficulties coded A LOT OF DIFFICULTY (3) or CANNOT DO AT ALL (4) in any domain, to 6: all six domains coded A LOT OF DIFFICULTY (3) or CANNOT DO AT ALL (4).

COUNT SUM_34 = VIS_SS HEAR_SS MOB_SS COM_SS COG_SS UB_SS (3 thru 4).
FREQUENCIES SUM_34.

SUM_34

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	.00	15454	89.2	89.2	89.2
	1.00	1367	7.9	7.9	97.1
	2.00	345	2.0	2.0	99.1
	3.00	117	.7	.7	99.8
	4.00	31	.2	.2	99.9
	5.00	9	.1	.1	100.0
	6.00	3	.0	.0	100.0
	Total	17326	100.0	100.0	

Step 5: Calculate Disability Identifier: DISABILITY1

The syntax below calculates the first disability identifier: **DISABILITY1** where the level of inclusion is at least one domain/question is coded SOME DIFFICULTY or A LOT OF DIFFICULTY or CANNOT DO AT ALL.
MISSING (9) are those who have coded 7, 8 or 9 on all six domains.

COMPUTE DISABILITY1 = 0.

IF (VIS_SS >= 7 and HEAR_SS >= 7 and MOB_SS >= 7 and COM_SS >= 7 and UB_SS >= 7 and COG_SS >= 7) DISABILITY1 = 9.

IF (SUM_234 >= 1) DISABILITY1 = 1.

NOTE: SUM_234 >= 1 means that at least one of the six domains is coded at least SOME DIFFICULTY (2).

> **VALUE LABELS** DISABILITY1 0 'without disability' 1 'with disability'.
> **RECODE** DISABILITY1 (9=SYSMIS).
> **FREQUENCIES** DISABILITY1.

DISABILITY1

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	without disability	9266	53.5	55.2	55.2
	with disability	7511	43.4	44.8	100.0
	Total	16777	96.8	100.0	
Missing		549	3.2		
Total		17326	100.0		

*Step 6: Calculate Disability Identifier: **DISABILITY2***

The syntax below calculates the second disability identifier: **DISABILITY2** where the level of inclusion is: at least 2 domains/questions are coded SOME DIFFICULTY or any 1 domain/question is coded A LOT OF DIFFICULTY or CANNOT DO AT ALL.
MISSING (9) are those who have coded 7, 8 or 9 on all six domains.

> **COMPUTE** DISABILITY2 = 0.
>
> **IF** (VIS_SS >= 7 and HEAR_SS >= 7 and MOB_SS >= 7 and COM_SS >= 7 and UB_SS >= 7 and COG_SS >= 7) DISABILITY2 = 9.
>
> **IF** (SUM_234 >= 2 OR SUM_34 = 1) DISABILITY2 = 1.

NOTE: The above syntax identifies those with at least two of the six domains coded as at least SOME DIFFICULTY (2): SUM_234 >= 2, OR those who have one domain that is coded A LOT OF DIFFICULTY (3) or CANNOT DO AT ALL (4): SUM_34 = 1.

> **VALUE LABELS** DISABILITY2 0 'without disability' 1 'with disability'.
> **RECODE** DISABILITY2 (9=SYSMIS).
> **FREQUENCIES** DISABILITY2.

DISABILITY2

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	without disability	12707	73.3	75.7	75.7
	with disability	4070	23.5	24.3	100.0
	Total	16777	96.8	100.0	
Missing		549	3.2		
Total		17326	100.0		

Step 7: Calculate Disability Identifier: DISABILITY3

The syntax below calculates the third disability identifier: **DISABILITY3** where the level of inclusion is: any 1 domain/question is coded A LOT OF DIFFICULTY or CANNOT DO AT ALL.
MISSING (9) are those who have coded 7, 8 or 9 on all six domains.
THIS IS THE CUT-OFF RECOMMENDED BY THE WG.

> **COMPUTE** DISABILITY3 = 0.
>
> **IF** (VIS_SS >= 7 and HEAR_SS >= 7 and MOB_SS >= 7 and COM_SS >= 7 and UB_SS >= 7 and COG_SS >= 7) DISABILITY3 = 9.
>
> **IF** ((VIS_SS = 3 or VIS_SS = 4) or (HEAR_SS = 3 or HEAR_SS = 4) or (MOB_SS = 3 or MOB_SS = 4) or (COM_SS = 3 or COM_SS = 4) or (UB_SS = 3 or UB_SS = 4) or (COG_SS = 3 or COG_SS = 4)) DISABILITY3 = 1.
>
> **VALUE LABELS** DISABILITY3 0 'without disability' 1 'with disability'.
> **RECODE** DISABILITY3 (9=SYSMIS).
> **FREQUENCIES** DISABILITY3.

DISABILITY3

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	without disability	14905	86.0	88.8	88.8
	with disability	1872	10.8	11.2	100.0
	Total	16777	96.8	100.0	
Missing		549	3.2		
Total		17326	100.0		

Step 8: Calculate Disability Identifier: **DISABILITY4**

The syntax below calculates the fourth disability identifier: **DISABILITY4** where the level of inclusion is any one domain is coded CANNOT DO AT ALL (4).
MISSING (9) are those who have coded 7, 8 or 9 on all six domains.

> **COMPUTE** DISABILITY4 = 0.
>
> **IF** (VIS_SS >= 7 and HEAR_SS >= 7 and MOB_SS >= 7 and COM_SS >= 7 and UB_SS >= 7 and COG_SS >= 7) DISABILITY4 = 9.
>
> **IF** ((VIS_SS = 4) or (HEAR_SS = 4) or (MOB_SS = 4) or (COM_SS = 4) or (UB_SS = 4) or (COG_SS = 4)) DISABILITY4 = 1.
>
> **VALUE LABELS** DISABILITY4 0 'without disability' 1 'with disability'.
> **RECODE** DISABILITY4 (9=SYSMIS).
> **FREQUENCIES** DISABILITY4.

DISABILITY4

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	without disability	16312	94.1	97.2	97.2
	with disability	465	2.7	2.8	100.0
	Total	16777	96.8	100.0	
Missing		549	3.2		
Total		17326	100.0		