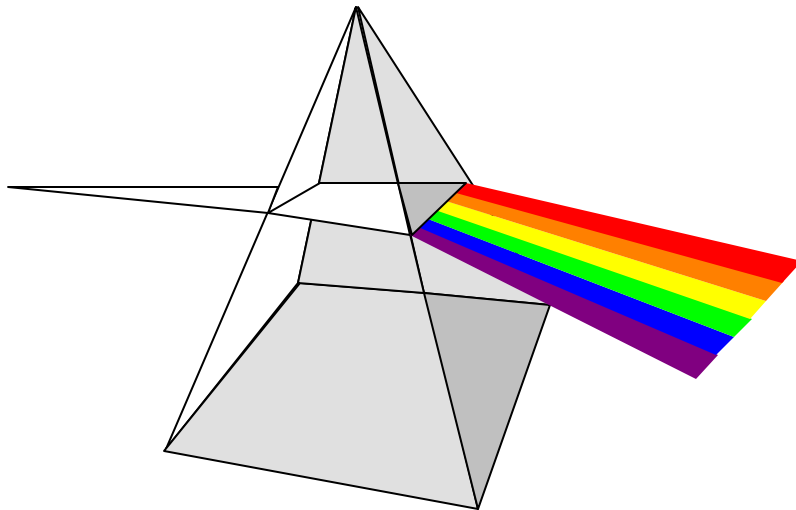


FACET Model



- Differences in raters' can account for as much variance as differences in examinee abilities.
- Assess rater severity, reliability, and adjust examinee scores for differences in raters.

FACET MODELS

FACET MODEL

A Facet model can be defined as:

$$\log [P_{nijx} / P_{nijx-1}] = \#_n - D_i - C_j - F_x$$

Where:

P_{nijx} = probability of person n being rated x on task i by judge j.

P_{nijx-1} = probability of person n being rated x - 1 on task i by judge j.

$\#_n$ = ability of person n

D_i = difficulty level of task i

C_j = severity of judge j

F_x = difficulty of rating step x-1 relative to x
(rating scale categories)

FACET MODELS

Chi-square Test (Ho: judge_i = judge_j, where i ... j)

$$X^2 = G (w_i D_i^2) - (G w_i D_i)^2 / G w_i$$
$$df = L - 1$$

Note: The value w_i , computed as $1/SE_i^2$, indicates the information for L measures, D_i , with standard errors, SE_i

A chi-square test examines the similarity among the facet elements, i. e., whether the L measures are statistically equivalent to one common "fixed" effect apart from measurement error. If $p > .05$, then L facet element measures are statistically different.

FACET EXAMPLE

RESEARCH FUNDING PROPOSALS

1. 60 research grant proposals received
2. 10 judges
3. Grant proposals rated on 1 - 5 scale (1= poor to 5 = excellent)
4. Each proposal received 2 ratings, but not by same 2 judges.

This is why Facets program is needed!

5. Proposals were ranked based on average of two ratings.
6. The top 10 grant proposals received funding.

Does which judge rating which proposal affect the outcome?

FACET EXAMPLE

FacForm Program

; File: faculty.key, converts two line per record file into a facform *.csv file

\$Input = faculty.dat ; flat file - ascii raw data

\$Output = faculty.fac ; facform comma separated file

\$Spoutput = faculty.spe ; specifications file

\$Facets=2 ; faculty and judge

\$Flabel=1,"proposal"

\$Flabel=2,"judge"

; Get items on the first line

\$DO=1

\$Label = 1,\$S1W3 ;proposal id in column 1-3

\$Label = 2,\$S5W2 ;first judge in column 5-6

\$Rating = \$S8W1 ;first rating in column 8

; Get items from the second line

\$Nextline

\$Label = 2,\$S5W2 ;second judge in column 5-6

\$Rating = \$S8W1 ;second rating in column 8

; Repeat for all subjects

\$AGAIN

FACET EXAMPLE

ASCII Data File (faculty.dat)

```
1 5 1
1 10 3
2 6 4
2 9 3
3 5 2
3 10 2
--
--
57 7 3
57 5 5
58 8 4
58 9 2
59 4 4
59 5 4
60 6 5
60 2 2
```

FacForm Program reads data as follows:

1st Line

Column 1-3 is Proposal ID

Column 5-6 is First Judge

Column 8 is Rating (1=poor to 5=excellent)

2nd Line

Column 1-3 is Proposal ID

Column 5-6 is Second Judge

Column 8 is Rating (1=poor to 5=excellent)

FACET EXAMPLE

FacForm Data File (faculty.fac)

1,5-10,1,,,,,3
2,6-9,4,,,3
3,5-10,2,,,,,2



57,5-7,5,,3
58,8-9,4,2
59,4-5,4,4
60,2-6,2,,,,5

Proposal 1 received a 1 from Judge 5 and
a 3 from Judge 10.



Proposal 60 received a 2 from Judge 2 and
a 5 from Judge 6.

FACET EXAMPLE

FacForm Specification File (faculty.spe)

; from file: FACULTY.KEY

Title= Facets Analysis of Peer Reviewed Grant Proposals

Facets = 2

Data file = faculty.fac

Scorefile = faculty

Output=faculty.out

Models =?,?,R5 ; ? = facet and R=# of ratings (1 to 5)

*

;Positive = 1

;Noncenter = 1

Labels =

1,proposal

1-60 ; 60 proposals

*

2,judge

1-10 ; 10 judges

*

FACET EXAMPLE

Facet Map of Proposals and Judges

Logit	proposal	judge	scale
+	4	+	+(5)

		*	
+	3	+	+
		*	---
+	2	+	+
		*	4

+	1	+	+ 10
		*	---

		***	1
		*	2 3 9
		*	5
*	0	*	* 3 *

		**	4
		*	6 8
		**	
		***	---
+	-1	+	+ 7
		**	

		*	2
		*	
+	-2	+	+
		*	---
		*	
+	-3	+	+(1)

Lenient Judges

Severe Judges

Chi-square= 106.4,

d.f. = 59, $p < .001$

Proposals were rated
significantly differently.

Reliability = .53

FACET EXAMPLE

Judge Rating Behavior

Judge	Total Ratings	Number of Ratings	Fair Average	Logit
1	29	9	2.7	.46
2	34	11	2.8	.33
3	37	12	2.8	.30
4	44	13	3.2	-.31
5	37	12	2.9	.20
6	46	13	3.4	-.60
7	62	16	3.6	-1.02
8	30	9	3.4	-.60
9	43	14	2.8	.26
10	31	11	2.4	.98

Chi-square= 21.6,
d.f.= 9, p = .01

Judges were
significantly
different in how
they rated
proposals.

Reliability = .52

Which judge(s) need training in rating grant proposals?

Judge 7

FACET EXAMPLE

TOP 10 PROPOSALS

Fair			
Proposal	Rating	Average	Logit
1	9	4.68	3.45
2	9	4.41	2.60
3	9	4.24	2.21
4	9	4.53	2.94
5	9	4.68	3.45
6	9	4.62	3.23
7	9	4.69	3.48
8	9	4.32	2.39
9	8	3.55	.89
10	8	3.80	1.32

FACET EXAMPLE

CONCLUSIONS

- ◆ The raw score rankings don't take into consideration the severity or leniency of the judge. Therefore, the fair average or logit estimate should be used to rank order the proposals before decisions are made regarding which proposals to fund.
- ◆ Some judges need to receive training or be excluded from rating proposals, i.e., Judge 7.
- ◆ Fair averages adjust raw score rankings for judge bias.

FACET ANALYSIS

SUMMARY

- ◆ **FACFORM** PROGRAM IS WRITTEN TO READ ASCII RAW DATA FILE AND OUTPUT TWO FILES: COMMA SEPARATED DATA FILE AND SPECIFICATION PROGRAM
- ◆ **FACET** PROGRAM IS RUN TO OBTAIN FAIR AVERAGES AND LOGIT ESTIMATES FOR FACET ELEMENTS
- ◆ **VERTICAL RULERS** ARE CREATED TO PERMIT VISUALLY INTERPRET FACET RELATIONSHIPS
- ◆ **CALIBRATED FILES** FOR EACH FACET CAN BE OUTPUT