


09/02/2017

The method of Comparative Judgement for assessment and research

And data analysis in R

San Verhavert
Sven De Maeyer
Vincent Donche



www.d-pac.be san.verhavert@uantwerpen.be

2

The method of Comparative Judgement for assessment and research


What?

- What is D-PAC?
- What is Comparative Judgement (CJ)?

How?

- Analysis in R
- Applications

- Research in D-PAC: in short



3

What is D-PAC NOT!




4

What is D-PAC?


- Digital Platform for the Assessment of Competences

1. Creating awareness
2. Developing a tool
3. Providing feedback



5


- What is D-PAC?
- What is Comparative Judgement (CJ)?
 - Terminology
 - Theoretical background
- Analysis in R
- Applications
- Research in D-PAC: in short



6

Terminology

- Comparative Judgement or CJ
- Assessment
- Competence
- Assessor, judge: individual who assesses
- Assessee: individual who is assessed
- Representation [of competence]: product that is assessed




7

Comparative Judgement (CJ)

- How do we assess?
- Assessment is difficult!
 - Consistent: over time, assessors, ...
- Making an absolute judgement is difficult and impossible (Laming, 1990)
- Every judgement is a comparison (Laming, 2003)

Rate the presenter:

7/20	20
16	9?
12	



8


Comparative Judgement

- Every judgement is a comparison (Laming, 2003)
- Thurstone (1927): The Law of Comparative Judgement

$$X_{AB} = v_A - v_B$$

$$p(A > B | v_a, v_b, \sigma_{ab}^2) = p(x_{AB} = 1 | v_a, v_b, \sigma_{ab}^2)$$

$$= \frac{1}{\sigma_{ij}\sqrt{2\pi}} \int_0^\infty \exp\left\{-\frac{(X - \mu)^2}{2\sigma^2}\right\} dt$$




9

Comparative Judgement

- Every judgement is a comparison (Laming, 2003)
- Thurstone (1927): The Law of Comparative Judgement
- Bradley-Terry-Luce model (BTL model) (Bradley & Terry, 1952; Luce, 1959)
- Rasch model (Rasch, 1960)

$$p(A > B | v_a, v_b) = p(x_{AB} = 1 | v_a, v_b) = \frac{\exp(S_A - S_B)}{1 + \exp(S_A - S_B)}$$


$$p(x_{vj} = 1 | \alpha_j, \tau_v) = \frac{\exp(\alpha_j - \tau_v)}{1 + \exp(\alpha_j - \tau_v)}$$




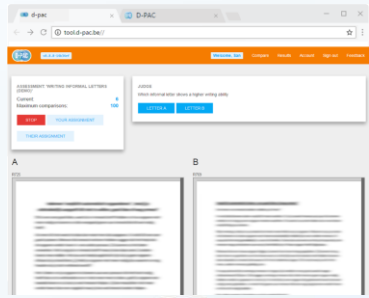
(Andrich, 1978)

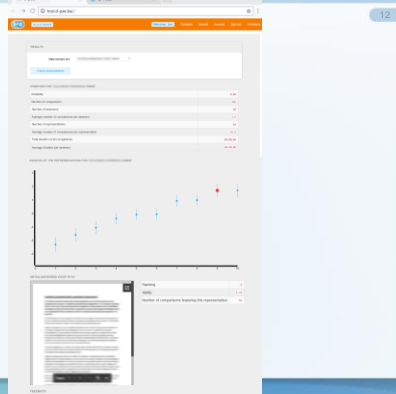
Comparative Judgement

- Writing and IQ testing: Thurstone, 1925; 1927
- Language assessment: Pollitt & Murray 1995
- Alternative assessment method (Pollitt, 2004; 2009)
- Facilitating evaluation of open-ended tasks (Jones, Swan & Pollitt, 2015)
- Inter-rater reliability $r = 0,86$ (Jones & Inglis, 2015)
- Efficiency




CJ now





13

- What is D-PAC?
- What is Comparative Judgement (CJ)?
- Analysis in R
 - The data structure
 - Estimating the BTL model in R
 - A simulation and estimation results
- Applications
- Research in D-PAC: in short




14

Analysis in R

comparison	assessment	assessor	representation A	representation B	selected representation	selected at	completed	Select base duration	Select base SD
3167	Papers	dpac11@inrets.be	Paper 1	Paper 12	Paper 12	10/12/2015 13:07	1	292	5
3169	Papers	dpac11@inrets.be	Paper 13	Paper 15	Paper 13	10/12/2015 13:15	1	178	6
3173	Papers	dpac11@inrets.be	Paper 22	Paper 8	Paper 22				
3177	Papers	dpac11@inrets.be	Paper 20	Paper 16	Paper 20				
3179	Papers	dpac11@inrets.be	Paper 5	Paper 9	Paper 5				
3183	Papers	dpac11@inrets.be	Paper 7	Paper 10	Paper 7				
3184	Papers	dpac11@inrets.be	Paper 3	Paper 10	Paper 3				
3192	Papers	dpac26@inrets.be	Paper 12	Paper 6	NA				
3185	Papers	dpac11@inrets.be	Paper 17	Paper 21	Paper 17				
3188	Papers	dpac23@inrets.be	Paper 2	Paper 14	Paper 2				
3188	Papers	dpac23@inrets.be	Paper 11	Paper 1	Paper 11				
3187	Papers	dpac11@inrets.be	Paper 6	Paper 20	Paper 6				
3190	Papers	dpac23@inrets.be	Paper 16	Paper 14	Paper 16				
3189	Papers	dpac23@inrets.be	Paper 7	Paper 20	Paper 7				
3020	Papers	dpac26@inrets.be	Paper 17	Paper 6	NA				
3191	Papers	dpac23@inrets.be	Paper 4	Paper 23	Paper 4				
3192	Papers	dpac23@inrets.be	Paper 12	Paper 13	Paper 12				
3194	Papers	dpac23@inrets.be	Paper 8	Paper 13	Paper 8				
3195	Papers	dpac23@inrets.be	Paper 11	Paper 13	Paper 11				
3199	Papers	dpac23@inrets.be	Paper 9	Paper 17	Paper 17	10/12/2015 16:35	1	293	7
3193	Papers	dpac32@inrets.be	Paper 5	Paper 3	Paper 5	10/12/2015 16:38	1	184	6

comparison number
assessment name
assessor e-mail address
representation A
representation B
selected representation
when selected
comparison completed
selection duration
selection perceived difficulty
...



15

```


72 ## calculate score ##
73 #int
74 score <- numeric(0)
75 for ( i in 1: length( data_int_raw[representation.A] ) )
76 {
77   if ( data_int_raw[representation.A][i] == data_int_raw[selected.representation(i) ] )
78     score <- append( score, 1 )
79   else if ( data_int_raw[representation.B][i] == data_int_raw[selected.representation(i) ] )
80     score <- append( score, 0 )
81   else
82     score <- append( score, na )
83 }
84 #int
85 #int
86 #int
87 #int
88 #int
89 #int
90 #int
91 #int
92 #int
93 #int
94 #int
95 #int
96 #int
97 #int
98 #int
99 #int
100 #int
101 #int
102 #int
103 #int
104 #int
105 #int
106 #int
107 #int
108 #int
109 #int
110 #int
111 #int
112 #int
113 #int
114 #int
115 #int
116 #int
117 #int
118 #int
119 #int
120 #int
121 #int
122 #int
123 #int
124 #int
125 #int
126 #int
127 #int
128 #int
129 #int
130 #int
131 #int
132 #int
133 #int
134 #int
135 #int
136 #int
137 #int
138 #int
139 #int
140 #int
141 #int
142 #int
143 #int
144 #int
145 #int
146 #int
147 #int
148 #int
149 #int
150 #int
151 #int
152 #int
153 #int
154 #int
155 #int
156 #int
157 #int
158 #int
159 #int
160 #int
161 #int
162 #int
163 #int
164 #int
165 #int
166 #int
167 #int
168 #int
169 #int
170 #int
171 #int
172 #int
173 #int
174 #int
175 #int
176 #int
177 #int
178 #int
179 #int
180 #int
181 #int
182 #int
183 #int
184 #int
185 #int
186 #int
187 #int
188 #int
189 #int
190 #int
191 #int
192 #int
193 #int
194 #int
195 #int
196 #int
197 #int
198 #int
199 #int
200 #int
201 #int
202 #int
203 #int
204 #int
205 #int
206 #int
207 #int
208 #int
209 #int
210 #int
211 #int
212 #int
213 #int
214 #int
215 #int
216 #int
217 #int
218 #int
219 #int
220 #int
221 #int
222 #int
223 #int
224 #int
225 #int
226 #int
227 #int
228 #int
229 #int
230 #int
231 #int
232 #int
233 #int
234 #int
235 #int
236 #int
237 #int
238 #int
239 #int
240 #int
241 #int
242 #int
243 #int
244 #int
245 #int
246 #int
247 #int
248 #int
249 #int
250 #int
251 #int
252 #int
253 #int
254 #int
255 #int
256 #int
257 #int
258 #int
259 #int
260 #int
261 #int
262 #int
263 #int
264 #int
265 #int
266 #int
267 #int
268 #int
269 #int
270 #int
271 #int
272 #int
273 #int
274 #int
275 #int
276 #int
277 #int
278 #int
279 #int
280 #int
281 #int
282 #int
283 #int
284 #int
285 #int
286 #int
287 #int
288 #int
289 #int
290 #int
291 #int
292 #int
293 #int
294 #int
295 #int
296 #int
297 #int
298 #int
299 #int
300 #int
301 #int
302 #int
303 #int
304 #int
305 #int
306 #int
307 #int
308 #int
309 #int
310 #int
311 #int
312 #int
313 #int
314 #int
315 #int
316 #int
317 #int
318 #int
319 #int
320 #int
321 #int
322 #int
323 #int
324 #int
325 #int
326 #int
327 #int
328 #int
329 #int
330 #int
331 #int
332 #int
333 #int
334 #int
335 #int
336 #int
337 #int
338 #int
339 #int
340 #int
341 #int
342 #int
343 #int
344 #int
345 #int
346 #int
347 #int
348 #int
349 #int
350 #int
351 #int
352 #int
353 #int
354 #int
355 #int
356 #int
357 #int
358 #int
359 #int
360 #int
361 #int
362 #int
363 #int
364 #int
365 #int
366 #int
367 #int
368 #int
369 #int
370 #int
371 #int
372 #int
373 #int
374 #int
375 #int
376 #int
377 #int
378 #int
379 #int
380 #int
381 #int
382 #int
383 #int
384 #int
385 #int
386 #int
387 #int
388 #int
389 #int
390 #int
391 #int
392 #int
393 #int
394 #int
395 #int
396 #int
397 #int
398 #int
399 #int
400 #int
401 #int
402 #int
403 #int
404 #int
405 #int
406 #int
407 #int
408 #int
409 #int
410 #int
411 #int
412 #int
413 #int
414 #int
415 #int
416 #int
417 #int
418 #int
419 #int
420 #int
421 #int
422 #int
423 #int
424 #int
425 #int
426 #int
427 #int
428 #int
429 #int
430 #int
431 #int
432 #int
433 #int
434 #int
435 #int
436 #int
437 #int
438 #int
439 #int
440 #int
441 #int
442 #int
443 #int
444 #int
445 #int
446 #int
447 #int
448 #int
449 #int
450 #int
451 #int
452 #int
453 #int
454 #int
455 #int
456 #int
457 #int
458 #int
459 #int
460 #int
461 #int
462 #int
463 #int
464 #int
465 #int
466 #int
467 #int
468 #int
469 #int
470 #int
471 #int
472 #int
473 #int
474 #int
475 #int
476 #int
477 #int
478 #int
479 #int
480 #int
481 #int
482 #int
483 #int
484 #int
485 #int
486 #int
487 #int
488 #int
489 #int
490 #int
491 #int
492 #int
493 #int
494 #int
495 #int
496 #int
497 #int
498 #int
499 #int
500 #int
501 #int
502 #int
503 #int
504 #int
505 #int
506 #int
507 #int
508 #int
509 #int
510 #int
511 #int
512 #int
513 #int
514 #int
515 #int
516 #int
517 #int
518 #int
519 #int
520 #int
521 #int
522 #int
523 #int
524 #int
525 #int
526 #int
527 #int
528 #int
529 #int
530 #int
531 #int
532 #int
533 #int
534 #int
535 #int
536 #int
537 #int
538 #int
539 #int
540 #int
541 #int
542 #int
543 #int
544 #int
545 #int
546 #int
547 #int
548 #int
549 #int
550 #int
551 #int
552 #int
553 #int
554 #int
555 #int
556 #int
557 #int
558 #int
559 #int
560 #int
561 #int
562 #int
563 #int
564 #int
565 #int
566 #int
567 #int
568 #int
569 #int
570 #int
571 #int
572 #int
573 #int
574 #int
575 #int
576 #int
577 #int
578 #int
579 #int
580 #int
581 #int
582 #int
583 #int
584 #int
585 #int
586 #int
587 #int
588 #int
589 #int
590 #int
591 #int
592 #int
593 #int
594 #int
595 #int
596 #int
597 #int
598 #int
599 #int
600 #int
601 #int
602 #int
603 #int
604 #int
605 #int
606 #int
607 #int
608 #int
609 #int
610 #int
611 #int
612 #int
613 #int
614 #int
615 #int
616 #int
617 #int
618 #int
619 #int
620 #int
621 #int
622 #int
623 #int
624 #int
625 #int
626 #int
627 #int
628 #int
629 #int
630 #int
631 #int
632 #int
633 #int
634 #int
635 #int
636 #int
637 #int
638 #int
639 #int
640 #int
641 #int
642 #int
643 #int
644 #int
645 #int
646 #int
647 #int
648 #int
649 #int
650 #int
651 #int
652 #int
653 #int
654 #int
655 #int
656 #int
657 #int
658 #int
659 #int
660 #int
661 #int
662 #int
663 #int
664 #int
665 #int
666 #int
667 #int
668 #int
669 #int
670 #int
671 #int
672 #int
673 #int
674 #int
675 #int
676 #int
677 #int
678 #int
679 #int
680 #int
681 #int
682 #int
683 #int
684 #int
685 #int
686 #int
687 #int
688 #int
689 #int
690 #int
691 #int
692 #int
693 #int
694 #int
695 #int
696 #int
697 #int
698 #int
699 #int
700 #int
701 #int
702 #int
703 #int
704 #int
705 #int
706 #int
707 #int
708 #int
709 #int
710 #int
711 #int
712 #int
713 #int
714 #int
715 #int
716 #int
717 #int
718 #int
719 #int
720 #int
721 #int
722 #int
723 #int
724 #int
725 #int
726 #int
727 #int
728 #int
729 #int
730 #int
731 #int
732 #int
733 #int
734 #int
735 #int
736 #int
737 #int
738 #int
739 #int
740 #int
741 #int
742 #int
743 #int
744 #int
745 #int
746 #int
747 #int
748 #int
749 #int
750 #int
751 #int
752 #int
753 #int
754 #int
755 #int
756 #int
757 #int
758 #int
759 #int
760 #int
761 #int
762 #int
763 #int
764 #int
765 #int
766 #int
767 #int
768 #int
769 #int
770 #int
771 #int
772 #int
773 #int
774 #int
775 #int
776 #int
777 #int
778 #int
779 #int
780 #int
781 #int
782 #int
783 #int
784 #int
785 #int
786 #int
787 #int
788 #int
789 #int
790 #int
791 #int
792 #int
793 #int
794 #int
795 #int
796 #int
797 #int
798 #int
799 #int
800 #int
801 #int
802 #int
803 #int
804 #int
805 #int
806 #int
807 #int
808 #int
809 #int
810 #int
811 #int
812 #int
813 #int
814 #int
815 #int
816 #int
817 #int
818 #int
819 #int
820 #int
821 #int
822 #int
823 #int
824 #int
825 #int
826 #int
827 #int
828 #int
829 #int
830 #int
831 #int
832 #int
833 #int
834 #int
835 #int
836 #int
837 #int
838 #int
839 #int
840 #int
841 #int
842 #int
843 #int
844 #int
845 #int
846 #int
847 #int
848 #int
849 #int
850 #int
851 #int
852 #int
853 #int
854 #int
855 #int
856 #int
857 #int
858 #int
859 #int
860 #int
861 #int
862 #int
863 #int
864 #int
865 #int
866 #int
867 #int
868 #int
869 #int
870 #int
871 #int
872 #int
873 #int
874 #int
875 #int
876 #int
877 #int
878 #int
879 #int
880 #int
881 #int
882 #int
883 #int
884 #int
885 #int
886 #int
887 #int
888 #int
889 #int
890 #int
891 #int
892 #int
893 #int
894 #int
895 #int
896 #int
897 #int
898 #int
899 #int
900 #int
901 #int
902 #int
903 #int
904 #int
905 #int
906 #int
907 #int
908 #int
909 #int
910 #int
911 #int
912 #int
913 #int
914 #int
915 #int
916 #int
917 #int
918 #int
919 #int
920 #int
921 #int
922 #int
923 #int
924 #int
925 #int
926 #int
927 #int
928 #int
929 #int
930 #int
931 #int
932 #int
933 #int
934 #int
935 #int
936 #int
937 #int
938 #int
939 #int
940 #int
941 #int
942 #int
943 #int
944 #int
945 #int
946 #int
947 #int
948 #int
949 #int
950 #int
951 #int
952 #int
953 #int
954 #int
955 #int
956 #int
957 #int
958 #int
959 #int
960 #int
961 #int
962 #int
963 #int
964 #int
965 #int
966 #int
967 #int
968 #int
969 #int
970 #int
971 #int
972 #int
973 #int
974 #int
975 #int
976 #int
977 #int
978 #int
979 #int
980 #int
981 #int
982 #int
983 #int
984 #int
985 #int
986 #int
987 #int
988 #int
989 #int
990 #int
991 #int
992 #int
993 #int
994 #int
995 #int
996 #int
997 #int
998 #int
999 #int
1000 #int

```

```

101 # int: "00-pac.be" off assessor name
102 data_ext$assessor <- sub(data_ext$assessor, pattern = "00-pac.be",
103 fixed = 1, replacement = "")
104

```



19

Estimating Bradley-Terry-Luce Model in R

- Our own function
 - CML
- btm function from sirt (Robitzsch, 2016)
 - Minorization-Maximization procedure (Hunter, 2004)

D-PAC

20

Estimating Bradley-Terry-Luce Model in R

- Our own function
 - CML
- btm function from sirt (Robitzsch, 2016)
- BTm function from BradleyTerry2 (Turner & Firth, 2012)
 - glm function
 - Least Squares Maximum Likelihood
- vglm and Brat functions from VGAM (Yee, 2010)
 - glm like function
 - Least Squares Maximum Likelihood

D-PAC

21

vglm from VGAM

vglm(Brat(DataMat) ~ 1,
 brat(refgp = refCatNo, refvalue = 1),
 trace = T)

	Paper 1	Paper 2	Paper 3	Paper 4
Paper 1	NA	5	1	0
Paper 2	2	NA	8	2
Paper 3	0	2	NA	10
Paper 4	10	8	0	NA


D-PAC

22

Estimating Bradley-Terry-Luce Model in R

- Our own function
 - CML
- btm function from sirt (Robitzsch, 2016)
- BTm function from BradleyTerry2 (Turner & Firth, 2012)
 - glm function
 - Least Squares Maximum Likelihood
- vglm and Brat functions from VGAM (Yee, 2010)
 - glm like function
 - Least Squares Maximum Likelihood

lbtPC.fit function in prefmod, btmodel function in psychotools, ...




23

Estimating Bradley-Terry-Luce Model in R

- Our own function
 - CML
- btm function from sirt (Robitzsch, 2016)
- BTm function from BradleyTerry2 (Turner & Firth, 2012)
 - Least Squares Maximum Likelihood
- vglm and brat functions from VGAM (Yee, 2010)
 - Least Squares Maximum Likelihood

Bias




24

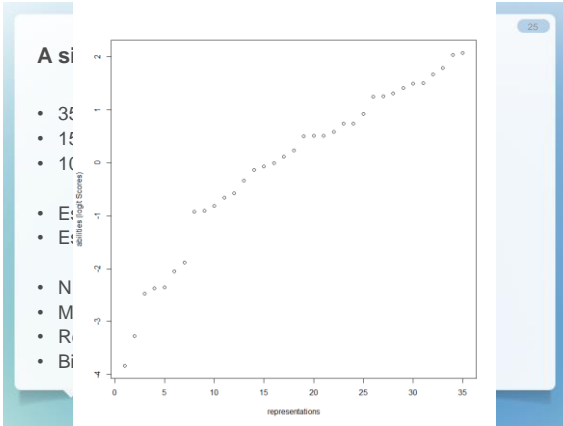
Estimating Bradley-Terry-Luce Model in R

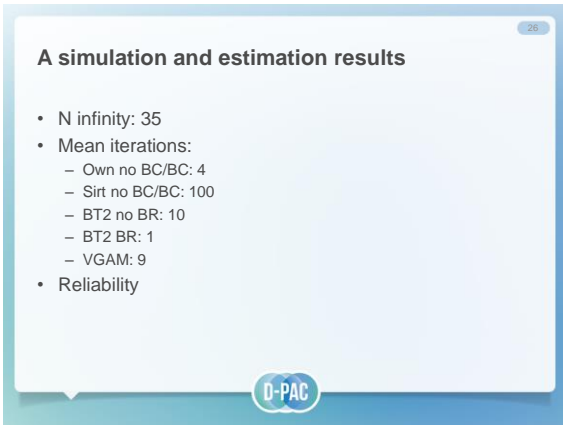
- Our own function
 - ϵ bias correction
- btm function from sirt (Robitzsch, 2016)
 - ϵ bias correction
- BTm function from BradleyTerry2 (Turner & Firth, 2012)
 - bias reduction by function modification
- vglm and brat functions from VGAM (Yee, 2010)

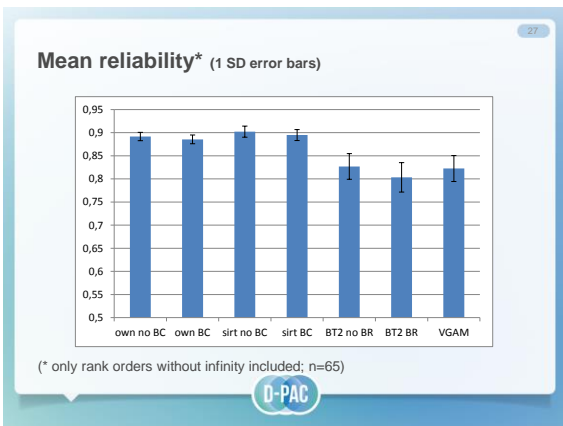
Bias
Unbound likelihood
Scale '0'

(Bertoli Barsotti & Punzo, 2012)









28

A simulation and estimation results

- N infinity: 35
- Mean iterations:
 - Own no BC/BC: 4
 - Sirt no BC/BC: 100
 - BT2 no BR: 10
 - BT2 BR: 1
 - VGAM: 9
- Reliability
- Bias

D-PAC

29

Mean bias (1 SD error bars)

Method	Mean Bias (approx.)	SE
own no BC	0.4	0.52
own BC	0.4	0.52
sirt no BC	0.6	0.54
sirt BC	0.4	0.54
BT2 no BR	0.8	0.69
BT2 BR	0.7	0.69
VGAM	0.75	?

D-PAC

30

A simulation and estimation results

- Conclusion
 - Our own function
 - btm from sirt with ϵ bias correction

D-PAC

31

- What is D-PAC?
- What is Comparative Judgement (CJ)?
- Analysis in R
- **Applications**
- Research in D-PAC: in short

D-PAC

32

Applications

Within the D-PAC project

Assessments

- # writing skills
- Reporting and reasoning
- ER models
- Moodboards

Jury's

- Selection based on rank order
- Consultation/meeting based on rank order
- Accountability (FB)

Research
26 try-outs

Assessment tool
10 try-outs

Learning tool
11 try-outs

Professionalization
1 try-out

Data collection and Scale development:

- Input PhD's
- # writing skills
- Reading skills
- Visual skills
- Mathematical problem solving
- Speech (audiologist)
- Number estimation
- Delphi study

Peer assessment

- Learning effect of comparing
- Learning effect of FB

(Maarten Goossens)

D-PAC

33


- What is D-PAC?
- What is Comparative Judgement (CJ)?
- Analysis in R
- Applications
- **Research in D-PAC: in short**

D-PAC

34

Research in D-PAC: in short

1. CJ method: reliability, efficiency and validity
2. Feedback to assessees
3. Feedback to organizations
4. Feedback to assessors
5. Design research: userfriendly?



Thank you!

San Verhavert
san.verhavert@uantwerpen.be

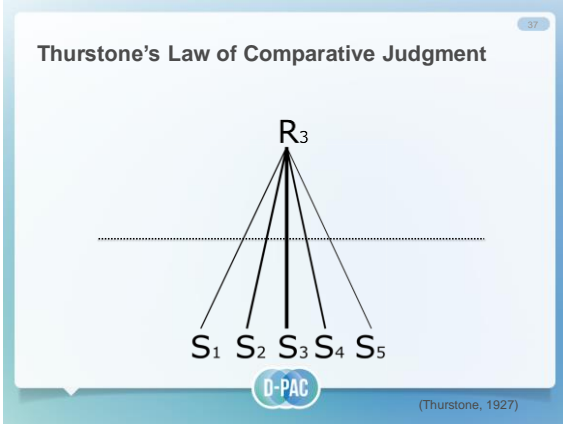


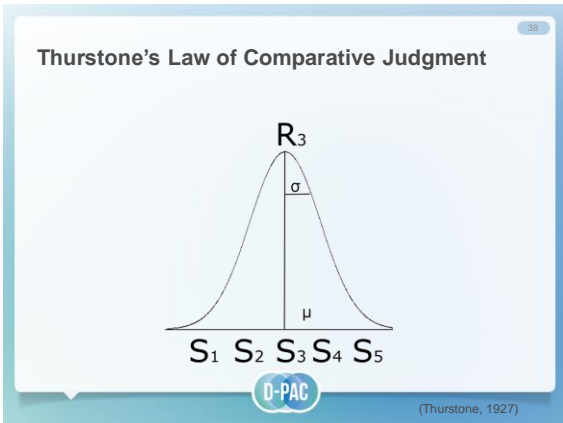


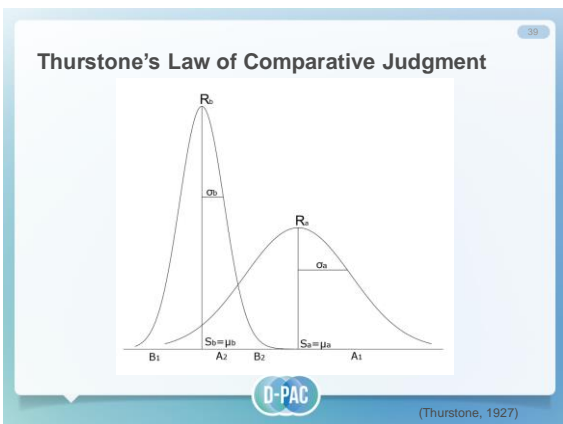
"Because human judgements are comparisons"

www.d-pac.be









40

Thurstone's Law of Comparative Judgment

$p(A > B | v_a, v_b, \sigma_{ab}^2) = p(x_{AB} = 1 | v_a, v_b, \sigma_{ab}^2)$
 $= \frac{1}{\sigma_{ij}\sqrt{2\pi}} \int_0^\infty \exp\left\{-\frac{(X-\mu)^2}{2\sigma^2}\right\} dt$

$X_{AB} = \frac{v_A - v_B}{\sigma}$
 $X_{AB} = v_A - v_B$

D-PAC

(Thurstone, 1927)

41

Bradley-Terry-Luce Model (BTL)

$p(A > B | v_a, v_b) = p(x_{AB} = 1 | v_a, v_b) = \frac{\exp(S_A - S_B)}{1 + \exp(S_A - S_B)}$
 $p(x_{vj} = 1 | \alpha_j, \tau_v) = \frac{\exp(\alpha_j - \tau_v)}{1 + \exp(\alpha_j - \tau_v)}$

$X_{AB} = v_A - v_B$

D-PAC

(Andrich, 1978)

42

Difference between methods

Method Comparison	Value (approx.)
own v sirt no BC	0.3
own v sirt BC	0.15
own v BT2 no BC/BR	0.6
own v BT2 BC/BR	0.4
own v vgam	0.55
sirt v BT2 no BC/BR	0.55
sirt v BT2 BC/BR	0.4
sirt v vgam	0.55
BT2 v vgam	0.05

D-PAC
