`mfcurve`: Visualizing Results From Multifactorial Designs

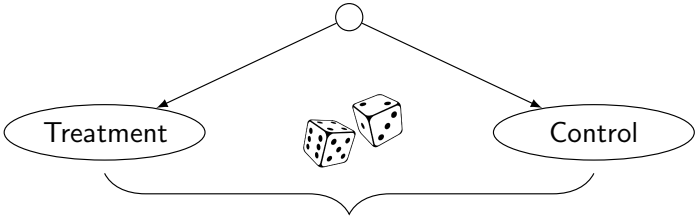
Daniel Krähmer
LMU Munich

2023 German Stata Conference
Berlin | 16 June

Funding by the DFG through the Priority Program META-REP (Project 464507200) is gratefully acknowledged.

The Classical Experimental Research Design

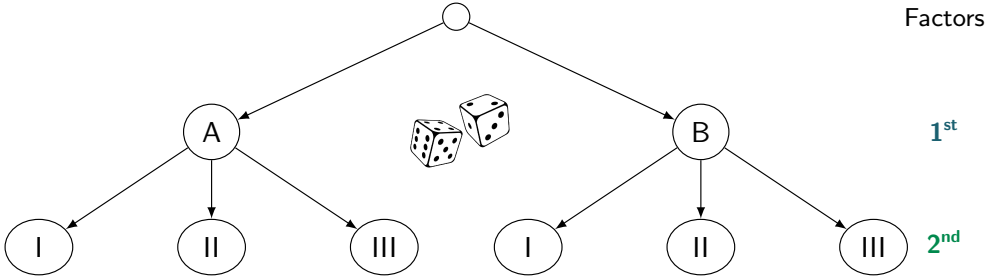
Estimand: Main effect of one particular treatment



$$ATE = E[Y^1] - E[Y^0]$$

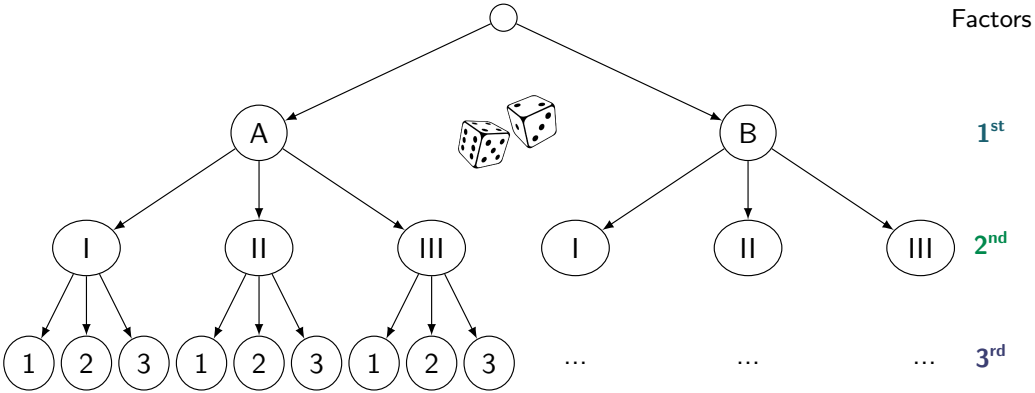
Multifactorial Research Designs

Estimand: Joint effect of several treatments



Multifactorial Research Designs

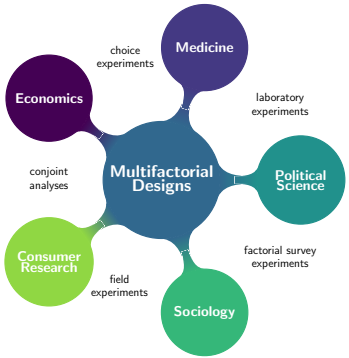
Estimand: Joint effect of several treatments



Multifactorial Research Designs Are...

... epistemologically useful,

... versatile & wide-spread,



Multifactorial Research Designs Are...

... epistemologically useful,

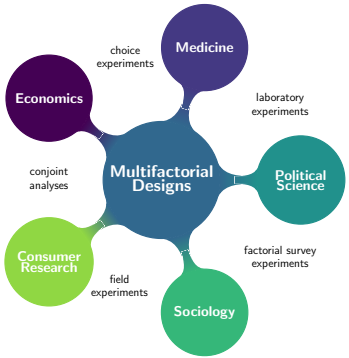

The NEW ENGLAND JOURNAL of MEDICINE

ORIGINAL ARTICLE

Hydroxychloroquine with or without Azithromycin in Mild-to-Moderate Covid-19

... versatile & wide-spread,

... **but difficult to visualize,** due to the spiraling number of treatment conditions
→ Potential remedy: `mfcurve`



Command `mfcurve`: Basics

`mfcurve` `depvar, factors(`*`indepvar`*`) [options]`

Input

- Outcome, i.e. dependent variable
- Factors, i.e. independent variables

▸ mechanics of `mfcurve`

Command `mfcurve`: Basics

`mfcurve` `depvar, factors(`*`indepvar`*`) [options]`

Input

- Outcome, i.e. dependent variable
- Factors, i.e. independent variables

Example: Average hourly wage across groups defined by race, region, and union membership (descriptive)

▸ mechanics of `mfcurve`

Command `mfcurve`: Basics

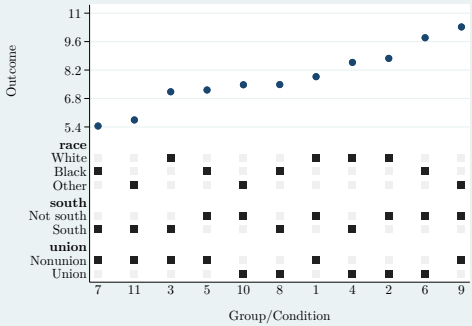
`mfcurve` `depvar, factors(`*`indepvar`*`) [options]`

Input

- Outcome, i.e. dependent variable
- Factors, i.e. independent variables

Example: Average hourly wage across groups defined by race, region, and union membership (descriptive)

```
sysuse nlsw88, clear
mfcurve wage, ///
    factors(race south union)
```



▸ mechanics of `mfcurve`

Command mfcurve: Options

Options	Description	
groupvar(*varname***)**	specify group identifier	
test(mean|zero)	perform significance tests	
level(#)	set confidence level	
show(*show_options***)**	add elements to the plot (see below)	
boxplot	use boxplots instead of point estimates	▸ demo
style...(*marker_options***)**	customize graph elements	▸ demo
twoway_options	twoway options, other than by()	

show_options	Description	
mean	add a horizontal mean line	
sig	highlight significant estimates	
ci_regular	add CIs, using solid lines	
ci_gradient	add CIs, using color gradients	▸ demo
groupsize	add case numbers to the x-axis	▸ demo

Command `mfcurve`: Application

Working Paper:
Care to Share? Experimental Evidence on Code Sharing Behavior in the Social Sciences

Command `mfcurve`: Application

Working Paper:
Care to Share? Experimental Evidence on Code Sharing Behavior in the Social Sciences

- What determines researchers' willingness to share analysis code upon request?
- Field experiment including more than 1,200 researchers across the social sciences
- Experimental variation of the code request's wording

Command `mfcurve`: Application

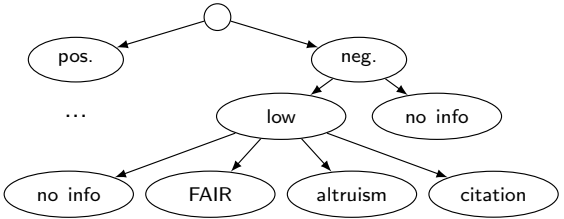
Working Paper:
Care to Share? Experimental Evidence on Code Sharing Behavior in the Social Sciences

- What determines researchers' willingness to share analysis code upon request?
- Field experiment including more than 1,200 researchers across the social sciences
- Experimental variation of the code request's wording

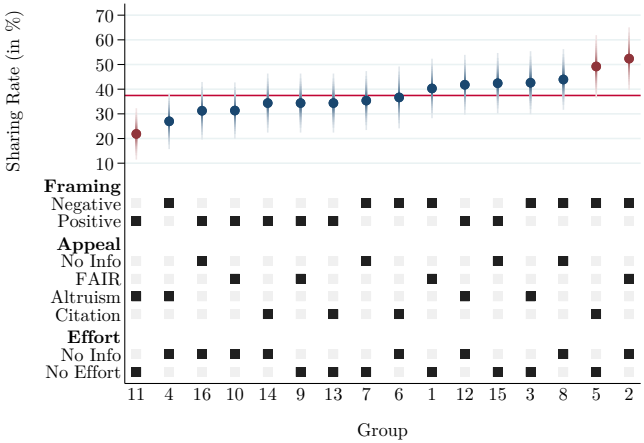
Framing (2 levels)

Effort (2 levels)

Appeal (4 levels)

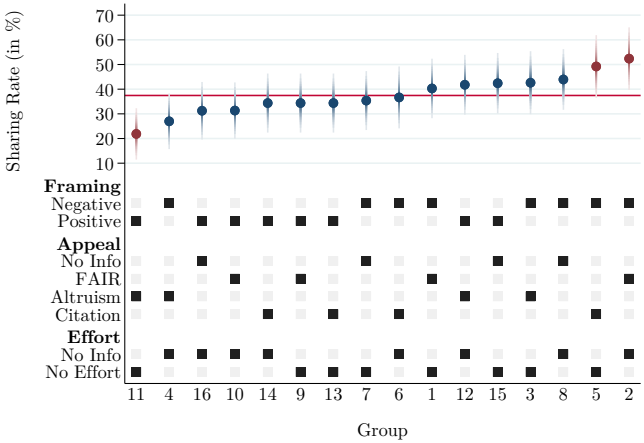


Command `mfcurve`: Application



▸ `mfcurve` vs. specification curves

Command `mfcurve`: Application

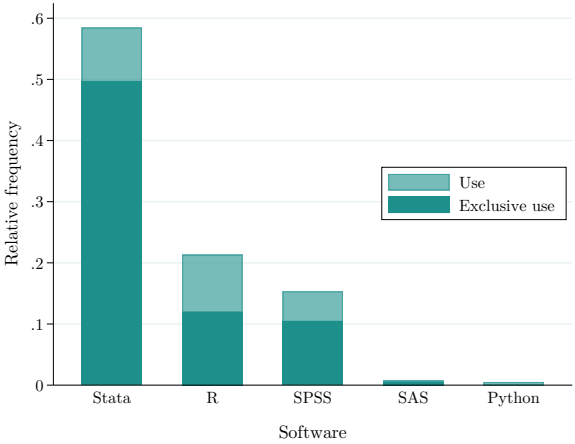


+ Transparent reporting of all results
+ Ranking of interventions by effectiveness
∼ Readability
∼ Main effects
∼ Interaction effects

▸ `mfcurve` vs. specification curves

By The Way

Software usage among authors who shared their code ($n = 385$)



Summary

- Multifactorial research designs are popular across disciplines
- They are notoriously difficult to visualize
- `mfcurve` may provide a solution to handle multidimensionality
- May also be used for simple *n*-dimensional description

- Installation from GitLab:
  ```
  net install mfcurve, from("https://tinyurl.com/mfcurve")
  ```

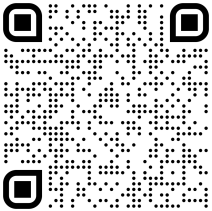
Thanks for your time!

Your comments and suggestions are appreciated.


@dkraehmer@sciences.social


daniel.kraehmer@soziologie.uni-muenchen.de

Project Homepage:



Mechanics of `mfcurve` I

Preparatory work:

race	south	union
Black	Not South	Union
Black	Not South	Nonunion
Black	South	Union
Black	South	Nonunion
White	Not South	Union
White	Not South	Nonunion
...	...	...

back

Mechanics of `mfcurve` I

Preparatory work:

- Define distinct groups in the n-dimensional space defined by variables in factor(...)

race	south	union	group
Black	Not South	Union	1
Black	Not South	Nonunion	2
Black	South	Union	3
Black	South	Nonunion	4
White	Not South	Union	5
White	Not South	Nonunion	6
...	...	...	...

back

Mechanics of `mfcurve` I

Preparatory work:

- Define distinct groups in the *n*-dimensional space defined by variables in factor(...)
- Calculate mean outcome by group

race	south	union	group	wage
Black	Not South	Union	1	9.79
Black	Not South	Nonunion	2	7.22
Black	South	Union	3	7.49
Black	South	Nonunion	4	5.45
White	Not South	Union	5	8.77
White	Not South	Nonunion	6	7.87
...	...	...	...	...

back

Mechanics of `mfcurve` I

Preparatory work:

- Define distinct groups in the *n*-dimensional space defined by variables in factor(...)
- Calculate mean outcome by group
- Rank groups by mean outcome

race	south	union	group	wage	rank
Black	Not South	Union	1	9.79	1
Black	Not South	Nonunion	2	7.22	6
Black	South	Union	3	7.49	4
Black	South	Nonunion	4	5.45	5
White	Not South	Union	5	8.77	2
White	Not South	Nonunion	6	7.87	3
...	...	...	...	...	...

back

Mechanics of `mfcurve` II

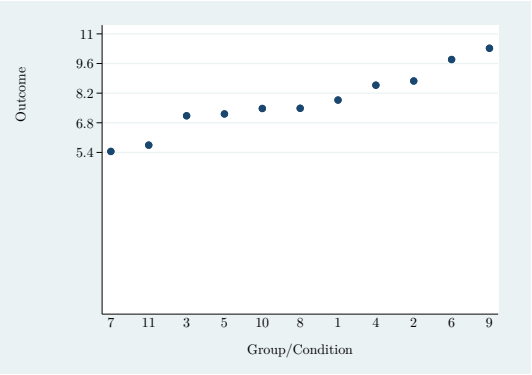
- Keep only one observation per group to increase efficiency (Stata Tip 19)
- Generate indicator variables, signaling each level's presence/absence

race	race_d_white	race_d_black	race_d_other	south	union	wage
Black	0	1	0	...	Union	9.79
Black	0	1	0	...	Nonunion	7.22
Black	0	1	0	...	Union	7.49
Black	0	1	0	...	Nonunion	5.45
White	1	0	0	...	Union	8.77
White	1	0	0	...	Nonunion	7.87
...	...	...	...	...	...	...

back

Mechanics of `mfcurve` III

Overlay plots:

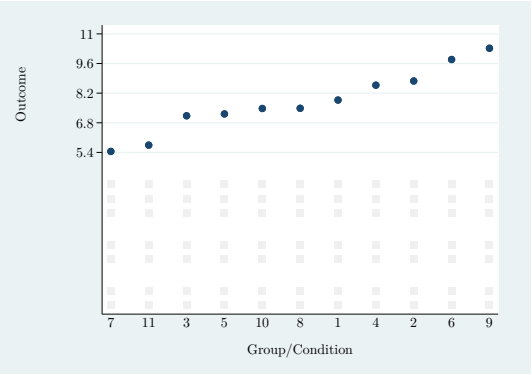


scatter outcome rank (using *group* labels!)

back

Mechanics of `mfcurve` III

Overlay plots:



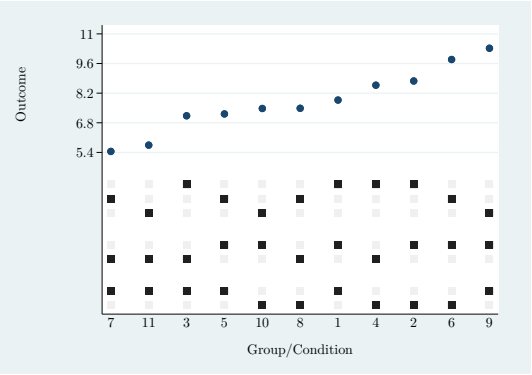
scatter outcome rank (using *group* labels!)

\+ indicators

back

Mechanics of `mfcurve` III

Overlay plots:



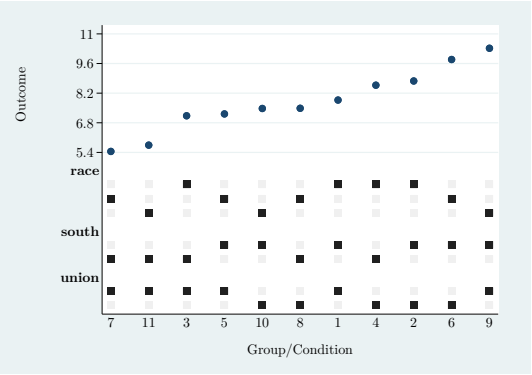
scatter outcome rank (using *group* labels!)

+ indicators
+ active indicators
  (based on dummies == 1)

back

Mechanics of `mfcurve` III

Overlay plots:



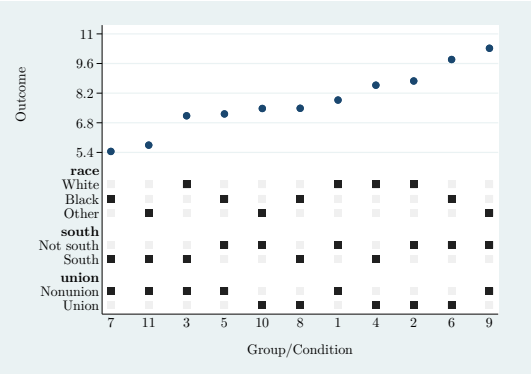
scatter outcome rank (using *group* labels!)

+ indicators
+ active indicators
  (based on dummies == 1)
+ ylabels based on *variable* labels

back

Mechanics of `mfcurve` III

Overlay plots:



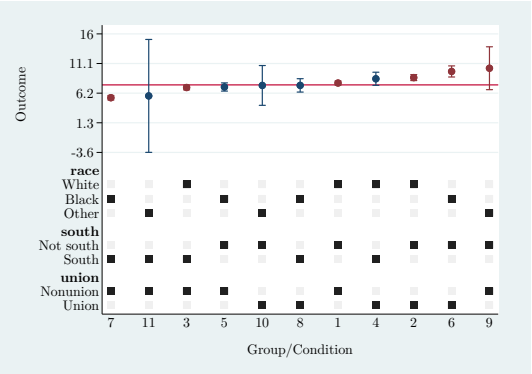
scatter outcome rank (using *group* labels!)

- \+ indicators
- \+ active indicators
  (based on dummies == 1)
- \+ ylabels based on *variable* labels
- \+ ylabels based on *value* labels

back

Mechanics of `mfcurve` III

Overlay plots:

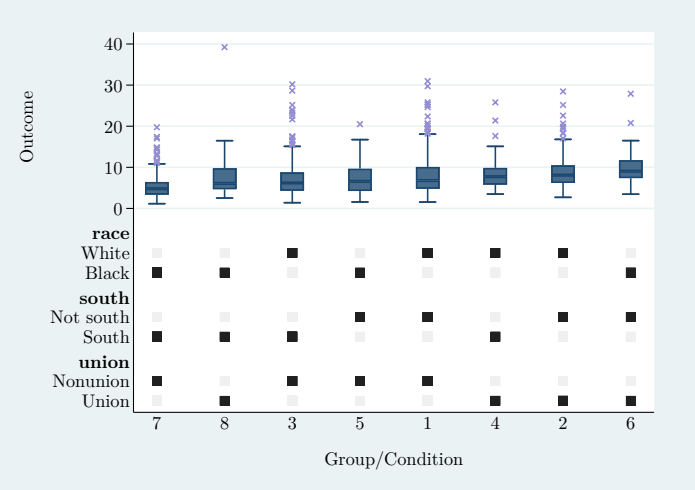


scatter outcome rank (using *group* labels!)

- \+ indicators
- \+ active indicators (based on dummies == 1)
- \+ ylabels based on *variable* labels
- \+ ylabels based on *value* labels
- \+ custom graph types (rcap, rspike, etc.)

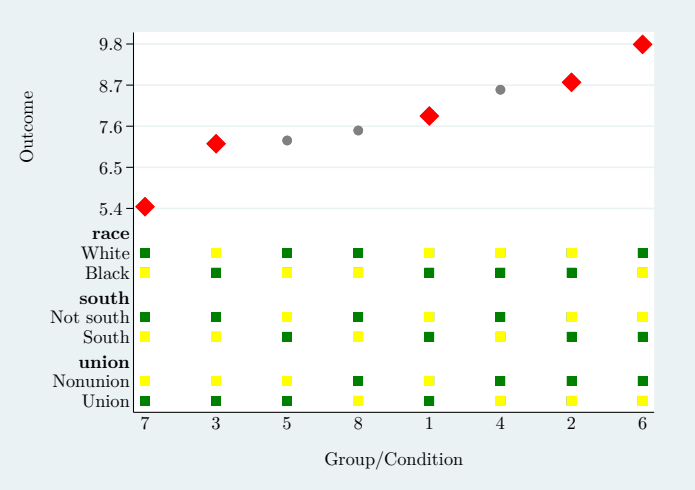
back

Demo: `boxplot`



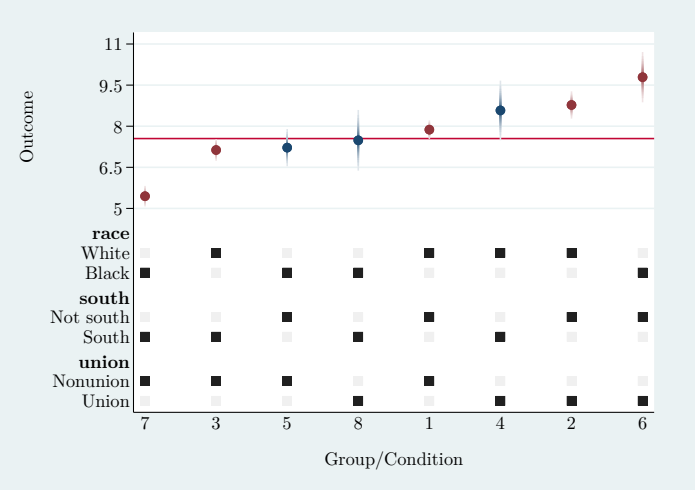
back

Demo: `style...(`*`marker_options`*`)`



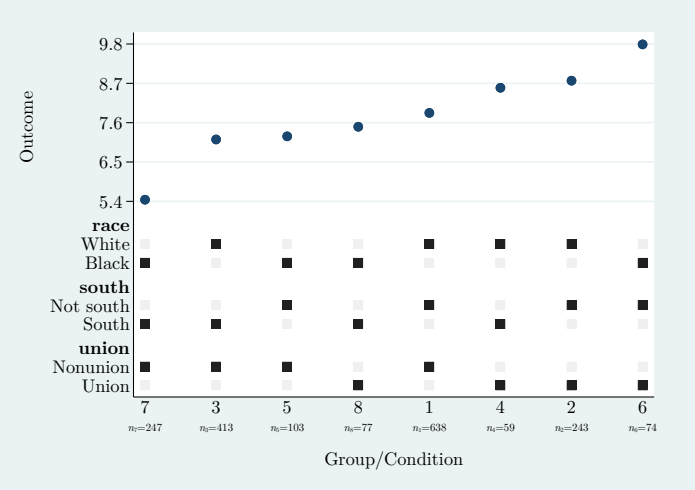
back

Demo: `show(mean ci_gradient)`



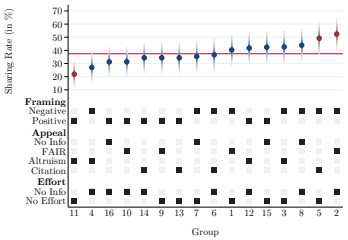
back

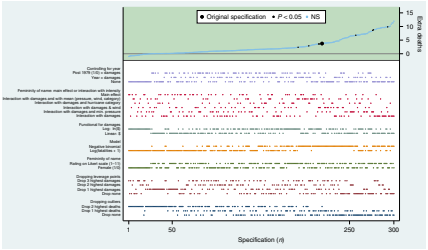
Demo: `show(groupsize)`



back

Comparison: `mfcurve` vs. specification curves



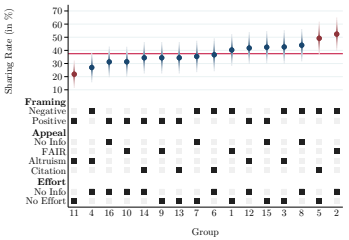


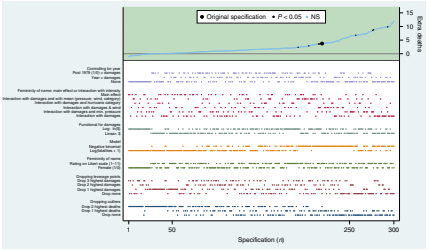
Graphical differences

upper panel:	mean values	multivariate regression coefficients
lower y-axis:	levels	model ingredients
x-axis:	treatment specifications	model specifications

back

Comparison: `mfcurve` vs. specification curves





Conceptual differences

analytical units:	discjunct subsamples	(overlapping) models
applications:	inferential & descriptive	inferential
computation:	parsimonious	intensive

back